CRÉDOC

CAHIER DE RECHERCHE

BILAN MÉTHODOLOGIQUE DES ÉVALUATIONS CONDUITES PAR LE CRÉDOC DANS LE CADRE DU FONDS D'EXPÉRIMENTATION POUR LA JEUNESSE

Isa ALDEGHI
Pauline JAUNEAU
Lara MULLER



DÉCEMBRE 2012

www.credoc.fr N° 300

SOMMAIRE

SYNTHESE 4

INTRODUCTION 8

PARTIE 1 : DE L'ELABORATION A LA MISE EN ŒUVRE D'UNE EXPERIMENTATION 15

I. LA GENESE DES PROJETS 15

I.1. Un premier appel à projets en avril 2009 pour améliorer l'insertion des jeunes 15

I.2. Les appels à projets suivants (2010 à 2013) 15

I.3. Les exigences méthodologiques de la Mission d'Animation du Fonds d'expérimentation pour la Jeunesse (MAFEJ) 17

I.4. Les porteurs de projet : une implication forte des missions locales 19

I.5. Les motivations des porteurs 19

I.6. Le contenu des projets 20

I.7. Une forte sollicitation du CRÉDOC pour évaluer des projets 21

I.8. Les propositions méthodologiques du CRÉDOC pour l'évaluation : une mixité de méthodes 21

I.9. Un temps très court pour rédiger les protocoles d'évaluation 23

II. LE LIEN ENTRE PORTEUR ET EVALUATEUR AU STADE DE LA REDACTION DU PROJET 24

II.1. L'évaluateur contribue à préciser les objectifs du projet 24

II.2. L'intervention des porteurs de projet sur les méthodes évaluatives : des raisons éthiques et pratiques 25

II.3. Associer un facilitateur 26

III. LA MISE EN ŒUVRE DES EXPERIMENTATIONS : DIFFICULTÉS RENCONTRÉES ET RECOMMANDATIONS 27

III.1. Une mobilisation du public parfois plus lente que prévue : la question de la montée en charge des dispositifs 27

III.2. Le développement des partenariats et des actions expérimentées plus long que prévu 29

III.3. Des recrutements temporaires parfois difficiles à effectuer 29

III.4. Pour des études de faisabilité préalables et bien cadrées 30

III.5. Des relations de confiance indispensables entre porteur de projet et évaluateur 31

PARTIE 2 : REGARD CRITIQUE SUR LES METHODES D'EVALUATION MISES EN ŒUVRE .. 32

I. DES APPROCHES QUANTITATIVES AMBITIEUSES A ADAPTER A LA TAILLE DES EXPÉRIMENTATIONS 32

I.1. Groupes bénéficiaire et témoin : l'idéal de la randomisation 33

I.1.1. Un tirage aléatoire en pratique souvent difficile à mettre en œuvre 33
I.1.2. Des solutions alternatives et leurs limites... 38
I.2. Les jeunes : un public peu accessible dans les enquêtes quantitatives...................... 41
I.2.1. Une forte attrition entre la première et la seconde interrogation 41
I.2.2. Les spécificités du public « jeunes » 42
I.2.3. Les enquêtes par questionnaire : cibler les indicateurs clés 43
I.3. Les tableaux de bord : des outils précieux qui nécessitent un remplissage harmonisé 44
II. DES APPROCHES QUALITATIVES A DEVELOPPER DANS LES EXPÉRIMENTATIONS 47
II.1. Les entretiens auprès des porteurs pour éclairer les conditions de mise en œuvre des actions.. 47
II.2. Les entretiens auprès des partenaires pour mettre en évidence les synergies et les freins 48
II.3. Les entretiens auprès des jeunes pour recueillir des récits de parcours et leur perception des dispositifs 49

PARTIE 3 : REFLEXIONS SUR LES CONDITIONS DE REUSSITE DES EXPERIMENTATIONS ET DE LEUR EVALUATION 51
I. LES CONDITIONS DE REUSSITE D'UNE EXPERIMENTATION 51
I.1. Une montée en charge anticipée pour évaluer les effets dès que les actions sont en place 51
I.2. Un public bien ciblé et clairement identifié 52
I.3. Des actions qui diffèrent clairement du droit commun 53
I.4. Des actions qui répondent à un besoin identifié et précis 54
I.5. Des financements qui coïncident avec la période d'évaluation 54
I.6. Un positionnement clair de l'évaluateur : une distance suffisante pour porter un jugement évaluatif 55
II. VALIDITE DES RESULTATS ET CONDITIONS DE GENERALISATION 56
II.1. Une durée minimale d'expérimentation pour des résultats à court et moyen terme 56
II.2. La validité externe : des effets reproductibles hors du cadre expérimental ? 57
II.2.1. Effets propres au caractère expérimental.. 57
II.2.1. Des interrogations sur l'extrapolation des résultats 59
II.2.2. La force des expérimentations... 61
III. USAGES ET UTILITE DES EVALUATIONS... 62
III.1. Usage et utilité des évaluations par les porteurs.. 62
III.2. Enseignements de politiques publiques 63

CONCLUSION.. 64

BIBLIOGRAPHIE 66

SYNTHESE

Dans un contexte national de développement des expérimentations sociales, le CRÉDOC a été amené à évaluer une vingtaine de dispositifs destinés aux jeunes de moins de 25 ans, dans le cadre du Fonds d'expérimentation pour la Jeunesse (FEJ). Le FEJ a financé depuis 2009 plus de 300 projets concernant les jeunes. Le présent Cahier de recherche dresse un bilan méthodologique des projets évalués par le CREDOC sur la période 2009-2013, qui ont porté sur l'insertion des jeunes sous main de justice, l'orientation de jeunes en rupture scolaire, l'accès à la formation en alternance, l'accès au logement et au microcrédit.

La première partie de ce Cahier de recherche retrace les étapes clés de l'élaboration des expérimentations dans le cadre du FEJ, de la genèse des projets à la conception de la méthodologie d'évaluation, qui devait si possible inclure, selon les préconisations du FEJ, une comparaison entre un groupe de bénéficiaires et un groupe témoin constitué aléatoirement (selon les principes d'une évaluation randomisée). Cette partie aborde l'implication des porteurs de projet, leurs motivations, ainsi que le contenu des projets suivis par le CRÉDOC. La deuxième partie de ce Cahier de recherche est consacrée aux aspects méthodologiques de l'évaluation et propose un tour d'horizon des méthodes quantitatives et qualitatives mobilisées par le CRÉDOC. Ce regard critique sur les méthodes mises en œuvre et les difficultés à les appliquer doit permettre de mieux définir les protocoles expérimentaux à l'avenir. La troisième partie du Cahier de recherche propose une réflexion sur les conditions d'obtention de résultats d'évaluation utiles pour la suite des expérimentations, en nous interrogeant d'une part sur la nature des résultats obtenus et, d'autre part, sur la portée et l'utilisation de ces résultats.

L'expérience acquise par le CREDOC dans le cadre du FEJ permet ainsi de formuler un certain nombre de recommandations sur les conditions de réussite des expérimentations et de leur évaluation.

Conditions d'émergence d'une expérimentation évaluable

Donner du temps pour rédiger les réponses à appel d'offres

Une des conditions à respecter pour que les expérimentations et leur évaluation puissent être menées à bien est de donner un temps suffisant aux porteurs de projet pour élaborer une proposition pertinente. Il est essentiel de préciser les objectifs du projet, de commencer à nouer des relations avec les partenaires-clé, d'évaluer la taille de la population cible de l'action et les moyens de l'atteindre et de la faire entrer dans l'expérimentation. Une fois tous ces éléments réunis, une durée minimale est nécessaire pour élaborer un protocole d'évaluation qui soit pertinent par rapport aux objectifs et aux actions de l'expérimentation. Un délai de six mois parait souhaitable au minimum pour assurer la cohérence de la méthode d'évaluation par rapport aux caractéristiques des projets évalués, dans le cadre d'une concertation entre porteurs et évaluateurs.

Identifier précisément les besoins et vérifier le caractère innovant de l'expérimentation

La question de la pertinence de l'action est à poser avant la mise en œuvre du projet. S'il apparait que les besoins tels qu'ils étaient présentés au départ ne correspondent pas à une demande du public, ou que des solutions existent déjà pour satisfaire les besoins, l'expérimentation ne présentera pas de plus-value identifiable. La définition d'un dispositif répondant à un besoin précis conditionne l'utilité de l'expérimentation et l'importance des effets qui peuvent être mesurés par l'évaluation. Si la différence entre le projet et le droit commun n'est pas assez marquée, l'impact de l'expérimentation ne pourra être que limité. Pour mesurer le caractère innovant d'un projet, l'évaluation doit ainsi comprendre une phase préalable d'analyse des actions et partenariats préexistants.

Avant le développement de l'expérimentation, prévoir une étape d'étude de faisabilité

Pour les projets d'une certaine taille, ainsi que pour ceux qui portent sur une population très éloignée de celle habituellement prise en compte par la structure porteuse de projet, il semble prudent de prévoir une étude de faisabilité qui estimera la durée réelle de montée en charge du dispositif, l'existence d'un besoin auquel répond l'action expérimentée, et la capacité des partenaires à mettre en œuvre l'action prévue. Cette phase doit se dérouler sur une période de trois à six mois pour que toutes ces dimensions puissent être explorées.

Faire intervenir un tiers acteur si un soutien au montage de projet est nécessaire

Dans le cadre de projets déployés sur plusieurs territoires, un soutien au montage de projet ou au pilotage peut être nécessaire pour assurer la cohérence globale et l'homogénéité de l'expérimentation sur les différents sites. Afin que l'évaluateur puisse conserver une position d'extériorité, il est parfois souhaitable d'associer un troisième acteur qui peut intervenir auprès du porteur de projet pour cette phase.

Forces et faiblesses des différentes méthodes d'évaluation

La taille critique des expérimentations pour conduire une évaluation quantitative

Les expérimentations de grande ampleur sont à privilégier si l'on souhaite mener une étude d'impact s'appuyant sur des méthodes quantitatives. Il est souvent nécessaire que l'action expérimentée concerne *a minima* plusieurs centaines de bénéficiaires pour disposer d'une puissance statistique suffisante pour détecter les effets de l'expérimentation. La précision obtenue dépendra à la fois de la taille des échantillons mais aussi de l'ampleur du phénomène analysé ou du niveau de l'indicateur mesuré.

Les modes de constitution d’un échantillon témoin

Le tirage aléatoire préconisé par le FEJ pour constituer le groupe de bénéficiaires et un échantillon témoin est rarement aisé à mettre en place. Des réticences de la part des porteurs de projets, des effectifs trop faibles de jeunes à suivre ou des problèmes de porosité des dispositifs évalués (impossibilité de réserver l’action expérimentée de façon exclusive au public cible) obligent parfois à abandonner cette méthode et à mettre en place d’autres types d’échantillons témoin.

Deux alternatives ont alors été mobilisées dans le cadre des projets évalués par le CRÉDOC et financés par la MAFEJ :

- un échantillon témoin suivi sur un autre territoire.
- un échantillon témoin constitué sur le même territoire à une période antérieure.

Ces méthodes, également limitées par la faiblesse des effectifs, n’ont pas non plus toujours permis d’identifier les effets propres au dispositif évalué, car les échantillons obtenus n’avaient pas les mêmes caractéristiques ou en raison de contextes économiques ou géographiques différents.

Les enquêtes par questionnaire à adapter au suivi des jeunes

Concernant les enquêtes quantitatives auprès des jeunes, celles-ci doivent tenir compte de l’attrition entre les différentes vagues. Les jeunes constituent un public avec lequel il est difficile de garder le contact. Le questionnement doit être court (15 minutes maximum) et ciblé sur les indicateurs attendus pour l’évaluation. Il doit contenir des questions simples et compréhensibles par tous pour appréhender les facteurs qui influent sur les parcours. Il ne doit en revanche pas chercher à appréhender trop de cas particuliers, nombreux au sein de la population des jeunes qui se caractérise par une grande hétérogénéité de situations.

Des tableaux de bord renseignés par les acteurs du programme

Les tableaux de bord doivent permettre de connaitre la situation des jeunes bénéficiaires et si possible également des témoins à différentes dates (a minima au début de l’expérimentation et à la fin). Ils comptabilisent aussi les actions suivies par les bénéficiaires et les témoins (droit commun et actions expérimentées). Ces outils de suivi sont, dans le cas où les données d’enquête font défaut, une source précieuse d’informations sur l’évolution des situations des jeunes. Les tableaux de bord doivent être aussi simples d’utilisation que possible pour les personnes en charge de leur remplissage, ciblés sur les indicateurs essentiels, et doivent également intégrer la nécessité pour l’évaluateur de pouvoir réaliser des analyses statistiques. Idéalement, quand il n’existe pas de logiciel de gestion, les tableaux de bord construits pour l’évaluation doivent comporter un masque de saisie pensé pour faciliter le traitement futur par l’évaluateur.

L'apport incontestable des enquêtes qualitatives

Les entretiens qualitatifs auprès des acteurs du programme ont apporté un regard complémentaire aux méthodes quantitatives pour bien comprendre les processus à l'œuvre, les réajustements et les évolutions dans le partenariat.

Quand les effectifs du public éligible à l'expérimentation sont trop faibles, les entretiens qualitatifs auprès de jeunes sont alors opportuns. Les entretiens approfondis auprès des jeunes ont permis d'aborder leur interprétation des effets des programmes sur leur parcours, leur ressenti du dispositif et leur compréhension des programmes.

Les clés d'une évaluation réussie

Au final, un certain nombre de grands principes peuvent être retenus au regard de ces expériences et du bilan réalisé par le CRÉDOC :

- Le délai de montée en charge des expérimentations doit être anticipé afin que le démarrage de l'évaluation coïncide au lancement effectif des projets.
- Le programme évalué doit se distinguer clairement des actions déjà en place sur le territoire (le droit commun) afin que l'impact des actions soient substantiel et donc identifiable.
- La validité externe des résultats (leur caractère reproductible dans le cadre d'un essaimage ou d'une généralisation) peut être limitée par les spécificités du contexte local mais aussi de la période de mise en œuvre du projet évalué.
- Les objectifs de l'évaluation doivent être cohérents avec la durée de la période d'observation. L'évaluateur doit disposer d'un recul suffisant pour mesurer l'effet d'un dispositif. L'expérimentation doit avoir une durée suffisante pour qu'il soit possible de mesurer tous les effets prévus.
- Les expérimentations sur plusieurs sites sont évaluables à condition de reposer sur des pratiques harmonisées grâce à un pilotage fort qui peut être assuré par un tiers.
- La sensibilisation des porteurs aux objectifs et exigences de l'évaluation est un atout pour assurer leur contribution aux recueils de données indispensables pour la démarche évaluative.

Au regard de ces enseignements, il est important de souligner que ces expérimentations ont créé un fort dynamisme tant chez les porteurs de projet que leurs partenaires, mais aussi chez les évaluateurs. On peut ainsi espérer qu'elles aient un effet catalyseur pour les politiques sociales, en ayant fait émerger des idées et des réflexions partagées dans le cadre d'une réflexion collective des différents acteurs locaux et nationaux. Il reste encore cependant à tirer et à croiser tous les enseignements des expérimentations financées par le FEJ, et à les mobiliser pour optimiser les dispositifs actuels, et les rendre à l'avenir plus efficaces et plus pertinents.

INTRODUCTION

Le recours aux expérimentations sociales se développe depuis le début des années 2000, en France comme dans d'autres pays européens. Rappelons-en la définition généralement retenue : « L'expérimentation dans le domaine des politiques sociales consiste à tester une innovation sociale à petite échelle et pendant une durée limitée, afin d'en mesurer les avantages et les inconvénients, de l'améliorer avant de la généraliser ou d'y renoncer si elle ne s'avère pas pertinente »[1].

Un cadre législatif français qui ouvre la voie aux expérimentations

En France, la démarche expérimentale est relativement récente dans l'histoire des politiques publiques soutenues par l'Etat, car elle se heurte à la tradition centralisée de l'Etat français ainsi qu'au principe de l'égalité entre les territoires. Le caractère pragmatique de la méthode « expérimenter, évaluer puis généraliser » peut s'accorder avec le mode de formation des décisions publiques dans les pays du « common law » tels que les Etats-Unis, la Grande Bretagne ou l'Irlande, où la jurisprudence est une source importante du droit[2]. En France, pays de droit romain où le système s'appuie davantage sur des codes écrits, et où les décisions sont prises, en principe, à la suite d'un débat public et d'une délibération entre les décideurs publics, les expérimentations sont entrées relativement récemment dans le paysage scientifique et politique, et ont été rendues possibles par les évolutions du contexte législatif et politique. C'est en effet la Révision constitutionnelle du 28 mars 2003 relative à l'organisation décentralisée de la République qui a inscrit le droit à l'expérimentation dans la loi[3], même si le Conseil Constitutionnel avait ouvert la voie dans un arrêt du 28 juillet 1993. L'article 37-1 de la Constitution dispose que « la loi et le règlement peuvent comporter, pour une durée et un objet limités, des dispositions à caractère expérimental ». L'article 72-4 permet aux collectivités locales d'adapter les lois et règlements nationaux aux situations locales. Elles peuvent donc être à l'initiative d'expérimentations dans leur domaine de compétences[4]. L'existence d'une étape d'évaluation n'est pas expressément mentionnée dans la Constitution, mais cette étape est inhérente à la notion même d'expérimentation. En effet, le principe de l'expérimentation est de tester l'hypothèse de l'effet d'un facteur sur un phénomène donné, et c'est l'évaluation qui permet de conclure à la validité ou non de l'hypothèse. Ainsi, les expérimentations sont associées au principe d'une évaluation indépendante qui permet de juger de l'intérêt des résultats obtenus avant de généraliser les politiques expérimentées. En fin de compte, la démarche est justifiée par ses fondements

---

[1] Cette définition, issue du colloque sur l'expérimentation sociale de Grenoble présidé par Martin Hirsch en novembre 2007, est reprise dans le guide méthodologique pour l'expérimentation sociale diffusé par la Mission d'Animation du Fonds d'Expérimentations pour le Jeunesse (MAFEJ) en 2009.

[2] Legendre, 2012.

[3] Loi constitutionnelle n° 2003-276 du 28 mars 2003 relative à l'organisation décentralisée de la République.

[4] Le droit de déroger est cependant subordonné à une habilitation préalable et doit avoir un objet et une durée limités.

scientifiques : dans un contexte de contrainte budgétaire, de besoin de modernisation des politiques sociales et de rationalisation des décisions publiques, l'expérimentation constitue un instrument offrant aux décideurs publics la possibilité de fonder leurs décisions ultérieures sur des données scientifiques à un coût raisonnable, puisqu'une expérimentation peut être déployée à petite échelle sans nécessairement couvrir l'ensemble du territoire national[5].

L'évaluation en tant que telle n'est pas pour autant nouvelle en France. Depuis la fin des années 80, l'impératif d'évaluation des politiques publiques est croissant et le principe de l'évaluation des lois émerge dans l'espace public français. L'évaluation commence à prendre une place plus centrale dans la conduite des politiques publiques avec la loi de décembre 1988 sur le Revenu Minimum d'Insertion (RMI). Votée à titre expérimental pour s'appliquer jusqu'en 1992, cette loi prévoyait la création d'une commission nationale d'évaluation chargée de réaliser un bilan des trois premières années d'application du RMI et d'émettre des propositions sur les conditions de son renouvellement[6]. Le cadre de l'évaluation des politiques publiques avait été fourni à cette occasion dans un rapport de Patrick Viveret au Premier Ministre[7], formalisé dans une circulaire portant sur « le renouveau du service public »[8] prônant l'évaluation comme un des axes à mettre en œuvre dans le cadre de la modernisation de l'État.

Les expérimentations contrôlées avec groupe témoin

Avec la révision constitutionnelle de 2003 et les orientations politiques qui s'ensuivent, ce qui est relativement nouveau c'est, d'une part, la possibilité d'expérimenter des dispositifs localement et temporairement, et, d'autre part, l'accent mis sur les évaluations d'impact incluant des comparaisons entre des groupes de bénéficiaires et des groupes témoins constitués idéalement par tirage au sort. Ces expérimentations avec groupe de contrôle tiré au sort (qu'on désigne aussi sous le terme d'expérimentations contrôlées avec randomisation) reposent sur la comparaison d'un groupe bénéficiant d'une mesure sociale avec un groupe témoin n'en bénéficiant pas, l'affectation des personnes éligibles à l'un des deux groupes de population se faisant aléatoirement afin de s'assurer que les deux groupes sont parfaitement semblables et que les différences observées sont attribuables uniquement à l'effet du programme expérimenté. Ce modèle d'innovation en matière de politiques sociales est utilisé depuis plusieurs décennies par les chercheurs des pays anglo-saxons[9]. Ses résultats ont été fortement diffusés et promus par le biais des grandes institutions internationales telles que la Banque Mondiale et des centres de recherches prestigieux comme le Massachusetts Institute of Technology (MIT) et son Laboratoire d'action contre la pauvreté ou J-Pal[10]. Créé en 2003 au sein du département d'économie du MIT par trois chercheurs en économie

---

[5] Chelle, 2012.

[6] Barbier, Maty-Jasik, 2010, p.125.

[7] Viveret, 1989.

[8] Circulaire du 23 février 1989 du premier Ministre (Michel Rocard).

[9] Dans le dernier quart du XXe siècle, environ 30 programmes basés sur une randomisation furent lancés aux États-Unis (Legendre, 2012).

[10] Abréviation de *Abdul Latif Jameel Poverty Action Lab*.

parmi lesquels la Française Esther Duflo, le J-Pal a fondé sa réputation sur l'utilisation exclusive d'expériences contrôlées pour mesurer les effets des programmes sociaux et de lutte contre la pauvreté dans les pays en développement. Mises en œuvre par les chercheurs de ses bureaux régionaux aux États-Unis, en Europe, en Afrique ou encore en Asie du Sud, ces méthodes s'exportent ainsi petit à petit en France sous la double influence du discours scientifique et politique dominant, affirmant la supériorité des méthodes d'expérimentations contrôlées pour mesurer l'impact de dispositifs, et pointant du doigt le « retard français ». De nombreux articles soulignent en effet l'entrée des expérimentations dans le champ des politiques publiques plus tardive en France que dans les pays anglo-saxons (Barbier, Maty-Jasik, 2010 ; Crépon 2008). Les conventions de tous les projets financés par le Fonds d'expérimentation pour la Jeunesse comportent ainsi en préambule le paragraphe suivant : « L'expérimentation sociale est en développement dans notre pays. Ce développement se réalise avec un certain retard par rapport aux démocraties anglo-saxonnes qui ont structuré depuis le milieu des années 1960 un important réseau de porteurs de projets expérimentaux et d'organisme d'évaluation. ».

Ces méthodes d'expérimentations contrôlées sont une transposition des essais cliniques réalisés dans le domaine de la médecine depuis les années 1920 pour démontrer l'efficacité d'un traitement, en comparant l'évolution de la situation d'un groupe de patients à qui est administré ce traitement, à celle d'un groupe témoin ne recevant aucun traitement, les deux groupes étant constitués par tirage au sort (Labrousse, 2010). Ces méthodes font cependant l'objet de critiques qui soulignent les limites de ces démarches : par exemple, des effets négatifs dus à l'interaction entre les traitements médicaux peuvent apparaître après la période de mise en œuvre de l'expérimentation[11]. La durée de l'observation peut être par ailleurs insuffisante pour voir émerger des effets secondaires qui peuvent se produire à plus long terme. Enfin, la validation d'un nouveau traitement peut se faire sur un résultat intermédiaire sans que les effets finaux visés ne soient démontrés. Ces critiques sont transposables aux expérimentations réalisées dans le champ des politiques sociales. Les améliorations constatées seront-elles durables ? Les résultats sur des objectifs intermédiaires (tels que le recours renforcé à des engrais par des agriculteurs) permettront-ils d'atteindre le résultat final (dans ce cas précis les aider à sortir de la pauvreté) ?

Aux États-Unis, après un développement à grande échelle dans les années 1960, les expérimentations avec groupes de contrôle aléatoire ont produit des résultats décevants, par manque de théorie explicative des comportements permettant d'analyser les causes des résultats obtenus[12]. La tendance est maintenant à la diversification des méthodes avec des évaluations davantage basées sur des théories d'action des politiques.

L'Ecole d'Economie de Paris, qui a vu le jour en 2006, a fait des expérimentations un de ses axes pédagogiques, avec la mise en place en 2007 d'un master Politiques Publiques et Développement qui vise à former des experts dans le domaine de l'évaluation de l'action publique, dans les pays en développement comme dans les pays développés. Deux projets à grande échelle ont été lancés sur cette période : l'expérimentation de l'accompagnement des chômeurs vers l'emploi par les

[11] Labrousse, 2010.

[12] Barbier, 2012.

opérateurs privés de placement, conduite en 2009 et 2010 par le Ministère du Travail, et l'expérimentation menée entre 2007 et 2009 en vue de la réforme du RMI et sa transformation en Revenu de Solidarité Active (RSA). La plus emblématique, l'expérimentation du RSA, n'est cependant pas la meilleure illustration de l'usage qui peut être fait de la démarche expérimentale : les territoires témoins dans les départements expérimentaux n'étaient pas strictement comparables aux territoires test. De plus, lancé tout d'abord dans plus d'une trentaine de départements volontaires, le RSA a en fait vu sa généralisation votée avant même que les conclusions définitives de l'évaluation ne soient publiées, cette décision ayant été prise en lien avec l'agenda politique[13].

Le cadre législatif national est par ailleurs en phase avec les orientations de l'Union Européenne (UE) en faveur de l'innovation locale qui se traduisent dans les interventions du Fonds Social Européen (FSE), et, depuis 2007, le programme de l'UE pour l'emploi et la solidarité sociale (PROGRESS), qui comporte explicitement parmi ses objectifs le soutien à l'expérimentation[14]. Ce programme considère que les expérimentations doivent être accompagnées d'un protocole d'évaluation afin de déterminer l'intérêt de leur généralisation. Parmi les méthodes d'évaluation décrites dans un guide méthodologique à l'intention des décideurs publics, les expérimentations randomisées sont présentées comme les plus fiables[15]. Plusieurs appels à projets en matière d'expérimentation sociale ont déjà été lancés dans le cadre de ce programme, avec des dotations allant jusqu'à 4,2 millions d'euros en 2013 pour soutenir environ cinq projets.

La création du Fonds d'expérimentation pour la Jeunesse

La nomination de Martin Hirsch comme Haut-Commissaire à la jeunesse le 12 janvier 2009 est le signe d'une volonté politique d'introduire l'expérimentation dans le champ des politiques sociales, l'ancien président d'Emmaüs étant un promoteur convaincu de ces démarches. Co-fondateur en 2005 de l'Agence nouvelle des solidarités actives (ANSA), qui a fait du développement des expérimentations son objet social, il a occupé le poste de Haut-Commissaire aux solidarités actives contre la pauvreté de 2007 à 2010. A ce titre, il avait présidé en novembre 2007 à Grenoble l'ouverture du « Grenelle de l'insertion » dont l'invité d'honneur était Gordon Berlin, le président d'un influent institut privé de recherches nord-américain qui travaille sur les expérimentations sociales depuis les années 1970[16]. C'est dans le cadre de sa mission aux solidarités actives

---

[13] Allègre, 2012.

[14] PROGRESS est un programme communautaire pour l'emploi et la solidarité sociale qui vise à financer des actions dans les domaines suivants : « préserver les emplois, en créer de nouveaux et stimuler la mobilité; renforcer les compétences et répondre aux besoins du marché du travail; et améliorer l'accès à l'emploi » (cf. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32010D0284:EN:NOT).

[15] *Social experimentation – A methodological guide for policy makers*, rédigé par J-Pal Europe à la demande de la Direction Générale de l'Emploi, des Affaires sociales et de l'Inclusion de la Commission européenne, septembre 2011. Cf. Jouen, 2012.

[16] Le *Manpower Demonstration Research Corporation*, créé en 1974 par le gouvernement des Etats-Unis et la fondation Ford. Pour plus de détails voir Chelle, 2012.

également qu'avait déjà été lancé, en 2009, un appel à projet d'expérimentations sociales (APES) visant à proposer de nouveaux outils pour réduire la pauvreté[17].

La mise en place, en 2009, du Fonds d'expérimentation pour la Jeunesse (FEJ) se situe donc dans la continuité de cette action volontariste et, de par l'ambition qu'elle se donne, marque un tournant significatif dans l'utilisation des expérimentations en France. Ce fonds d'appui aux expérimentations en faveur des jeunes, institué par l'article 25 de la loi du 1er décembre 2008[18], a pour objet de financer des programmes expérimentaux visant à favoriser la réussite scolaire des élèves, à contribuer à l'égalité des chances et à améliorer l'insertion sociale et professionnelle des jeunes de moins de vingt-six ans[19]. Animé par l'Etat (Direction de la jeunesse, de l'éducation populaire et de la vie associative), sa gestion est confiée à la Caisse des Dépôts et Consignations. Un conseil scientifique, composé de dix personnalités qualifiées dans le domaine de l'évaluation, est chargé de garantir la qualité des évaluations mises en œuvre. En 2011, les attributions du fonds sont élargies : il peut en outre financer « l'essaimage d'expérimentations dans de nouveaux territoires »[20]. Doté de près de 230 millions d'euros[21], le fonds vise à susciter des expérimentations en faveur de l'insertion des jeunes et fonctionne principalement sur le mode d'appels à projets portant sur des thématiques ciblées.

Le fonds, qui a pour ambition de créer et de financer des politiques innovantes de la jeunesse, doit ainsi permettre d'impulser et de soutenir des initiatives, sur différents territoires et selon des modalités variables - et de les évaluer. Cette approche vise ainsi à préparer, dès la phase expérimentale, les conditions d'une généralisation ultérieure. Le fonds a conduit à la mise en œuvre au niveau local de plus de 500 projets. Sa particularité est d'insister sur la supériorité des méthodes randomisées, dont l'utilisation est fortement préconisée dans le guide méthodologique conçu pour les porteurs de projets[22].

---

[17] Au total, 37 projets avaient été sélectionnés. Un budget de 9 millions d'euros a été dégagé pour financer ces projets en 2008 et en 2009. Le CREDOC en a évalué trois entre 2008 et 2012.

[18] La loi 2008-1249 du 1er décembre 2008 généralisant le Revenu de Solidarité Active et réformant les politiques d'insertion.

[19] L'appel à projets est structuré autour de 14 thématiques : alternance, jeunes sous mains de justice, décrochage scolaire, décrochage universitaire, engagement des jeunes, insertion professionnelle, livret de compétences, logement, lutte contre les discriminations, microcrédit, mobilité internationale, orientation, réussite scolaire, santé.

[20] Extrait du décret n° 2011-1603 du 21 novembre 2011 relatif au fonds d'appui aux expérimentations en faveur des jeunes.

[21] 170 millions de subventions de l'Etat (crédits de la jeunesse et du Plan de relance) et 55 millions provenant de partenaires privés, dont 50 millions de la fondation d'entreprise Total. Cf. http://www.jeunes.gouv.fr/ministere-1001/actualites-1016/zoom-sur-800/article/le-fonds-d-experimentation-pour-la

[22] Guide méthodologique de la MAFEJ pour l'évaluation des expérimentations sociales.

Les évaluations d'expérimentations menées par le CRÉDOC

Dans le cadre du premier appel à projets du FEJ, le CRÉDOC a été fortement sollicité pour être évaluateur. Entre 2009 et 2012, le département Evaluation des Politiques Sociales (EPS) du CRÉDOC a évalué seize expérimentations en faveur des jeunes[23] et le département Evaluation des Politiques Publiques en a évalué deux[24], dans le cadre des différents appels à projets du FEJ.

Le présent Cahier de recherche a pour objectif de dresser le bilan méthodologique de ces évaluations pour le Fonds d'Expérimentation pour la Jeunesse dont la plupart se sont achevées en 2012. Cette recherche s'appuie également sur une expérience en matière d'évaluation acquise au cours d'autres expérimentations conduites récemment. Le CRÉDOC a notamment contribué à l'évaluation du RSA en 2010-2011 (*via* une enquête qualitative auprès de 360 allocataires)ainsi que du projet des Lits d'Accueil Médicalisés entre 2009 et 2012 pour la Direction Générale de la Cohésion Sociale (DGCS) pour citer les principales. L'évaluation d'une expérimentation sur le Revenu Contractualisé d'Autonomie (RCA)[25] pour le Ministère de la Jeunesse est par ailleurs en cours actuellement.

Un bilan méthodologique à destination des porteurs de projet et des évaluateurs

Le présent travail fait suite à un précédent Cahier de recherche du CRÉDOC qui dressait en 2009 les tout premiers constats sur les conditions d'élaboration et de démarrage des premiers projets évalués par le département EPS portant sur les politiques locales en faveur de la jeunesse[26].

Il s'agit ici de mettre en évidence les bonnes pratiques et les écueils liés aux difficultés, pratiques et théoriques, qui sont apparues tout au long des expérimentations, de la conception à la mise en œuvre concrète des expérimentations et de l'évaluation. L'analyse critique des processus et méthodes proposée ici vise à tirer les enseignements utiles à de futures évaluations. Trois ans après le lancement des premiers projets financés par le FEJ, les retours d'expérience se multiplient[27].

Dans un premier temps, nous retraçons les étapes clés de l'élaboration d'une expérimentation sociale, de la genèse des projets à la conception d'une méthodologie adaptée en nous appuyant sur l'expérience du CRÉDOC. Nous retraçons l'historique des projets lancés par le Fonds d'expérimentation pour la Jeunesse, avec ses exigences méthodologiques, l'implication des porteurs de projet et les interactions entre évaluateur et porteurs de projet.

La deuxième partie est consacrée aux aspects méthodologiques de l'évaluation et propose un tour d'horizon des méthodes quantitatives et qualitatives mobilisées. Ce regard critique sur les

---

[23] Trois évaluations ont été réalisées en collaboration avec des co-évaluateurs, respectivement le Centre de recherche en économie et en statistiques (CREST), le cabinet MB-Consultants, et l'espace Inter Initiatives - E2I.

[24] Trois évaluations d'expérimentations pour la jeunesse étaient encore en cours fin 2012 au CRÉDOC.

[25] Le CREST est co-évaluateur.

[26] Brézault, Olm, Angotti, Simon, 2009.

[27] Parmi ces analyses, on peut mentionner : Divay 2011 et 2012, Gilles 2012, Gurgand et Valdenaire 2012, ainsi que le Rapport du Conseil scientifique du Fonds d'expérimentation pour la jeunesse pour l'année 2011.

méthodes mises en œuvre et les difficultés à les appliquer est destiné à permettre de mieux définir les protocoles expérimentaux à l'avenir.

Enfin, la troisième partie est consacrée aux conditions d'obtention de résultats d'évaluation utiles pour la suite des expérimentations, en nous interrogeant d'une part sur la nature des résultats obtenus et, d'autre part, sur la portée et l'utilisation de ces résultats.

PARTIE 1 : DE L'ELABORATION A LA MISE EN ŒUVRE D'UNE EXPERIMENTATION

I. LA GENESE DES PROJETS

I.1. Un premier appel à projets en avril 2009 pour améliorer l'insertion des jeunes

Le Fonds d'expérimentation pour la Jeunesse (FEJ) a lancé une première vague d'appel à projets (AP1) le 6 avril 2009 avec pour axes principaux :

- La réduction des sorties prématurées du système de formation initiale.
- L'amélioration de la transition entre formation et emploi.
- La prévention des ruptures : expérimentations dans le domaine du logement, actions auprès des jeunes sortants de l'ASE-PJJ, de jeunes sous main de justice et intervention auprès d'étudiants en difficulté.
- Le soutien des initiatives et des projets de jeunes (microcrédit et mobilité internationale).

Le budget consacré au financement de ces projets était de 55 millions d'euros. Il pouvait s'agir de projets concernant la France métropolitaine comme les territoires et départements d'outre-mer.

Pour pouvoir prétendre à un financement par le Fonds d'expérimentation pour la Jeunesse, chaque projet devait prévoir son évaluation par une équipe extérieure. Cette évaluation était obligatoire, aucun projet ne pouvant être accepté sans qu'un protocole évaluatif n'y soit associé. Tandis que l'évaluation pouvait être financée intégralement par le FEJ, les projets expérimentaux devaient toujours faire l'objet d'un cofinancement avec un ou plusieurs partenaires, et le FEJ ne pouvait contribuer au plus qu'à 50% du budget global.

En 2009, le Fonds d'expérimentation pour la Jeunesse a reçu près de 500 propositions et en a retenu environ un tiers[28].

I.2. Les appels à projets suivants (2010 à 2013)

Trois autres appels à projets concernant les jeunes ont été lancés entre mai 2010 et février 2011 (AP2, AP3, AP4). L'AP2 est centré sur le thème de la lutte contre les discriminations. L'AP3 couvre l'engagement des jeunes et l'élargissement des choix professionnels des filles. L'AP4 porte sur les pratiques culturelles. Le nombre de projets présentés et acceptés est nettement plus réduit que pour l'AP1 (43 pour l'AP2, 24 pour l'AP3 et 6 pour l'AP4). Le CRÉDOC a été associé à l'évaluation

---

[28] Sur 494 propositions, 169 ont été retenues.

d’un projet de l’AP2 sur la discrimination et d’un projet de l’AP3 sur l’élargissement des choix d’orientation des jeunes filles.

Tableau 1 : Expérimentations pour la Jeunesse évaluées par le CRÉDOC

Code MAFEJ	Nom du projet	Porteur
Thème	**Jeunes sous main de justice**	
AP1_041	Réussir sa sortie (R2S) (jeunes sous main de justice)	Mission locale (ML) des Ulis (avec la participation des ML de l’Essonne)
AP1_067	Action Pour la REinsertion Sociale et professionnelle des personnes placées sous main de justice (APRES Jeunes)	Association *Entrepreneurs et Développement*
AP1_348	Un pas dehors (jeunes sous main de justice)	ML Le Mans
AP1_457	Accompagner Vers l’Emploi et Construire un prcours personnalisé (AVEC)	ML Toulouse et ML de Haute-Garonne
Thème	**Orientation / décrochage**	
AP1_173	Osons l'égalité	Chambre de Commerce et d’Industrie 22
AP1_274	Orientation active	Académie d'Aix-Marseille
AP1_299	De l'université à l'entreprise	Association *Entreprendre ensemble* (gérant la ML de Dunkerque)
AP1_492	Plateformes décrocheurs	DASES – Conseil Général (CG) de Paris
Thème	**Alternance**	
AP1_037	DEPAR - Dispositif expérimental de plateforme pour une alternance réussie	Fondation Apprentis d'Auteuil
AP1_112	Promotion de l'apprentissage	CNML (coordination)
AP1_329	Champagne Alternance	Association Régionale des Présidents de ML
Thème	**Logement**	
AP1_079	Plateforme Logement Jeunes 94	CG 94
AP1_195	Développement d’un service Médiation logement	ML d’Agen, ML de Villeneuve-sur-Lot et ML de Marmande
AP1_269	Agiloje-Agir pour le Logement des Jeunes	CNML (coordination), UNHAJ, UNCLLAJ, Fédération des PACT, FAPIL
Thème	**Microcrédit**	
AP1_088	Microcrédit	Centre Communal d’Action Sociale (CCAS) de Grenoble
AP1_411	Microcrédit	CCAS de Besançon
Thème	**Lutte contre la discrimination**	
AP2_086	Une chance pour tous …De nouvelles approches du recrutement pour de nouveaux talents en entreprise	Association Régionale des Présidents de ML Limousin
Thème	**Diversification des choix d’orientation scolaire et professionnelle des jeunes filles**	
AP3_048	Futures scientifiques	Association *Lunes et l’autre* (avec la participation du Rectorat de la région PACA)

ML : Mission locale – CG : Conseil général

Entre avril 2010 et janvier 2013, trois autres appels à projets ont été lancés (APDOM1 à APDOM3). Consacrés aux départements d’Outre-Mer, ils concernent chacun entre 11 et 15 projets. De plus, quatre appels à projets, lancés entre août 2011 et avril 2012, étaient plus spécifiquement liés au milieu scolaire (APSCO1 à APSCO4). Ils portent chacun sur une thématique précise : les rythmes scolaires, l’accompagnement des bacheliers professionnels avant et après l’examen final, le

harcèlement en milieu scolaire et le développement de l'« esprit d'entreprendre » auprès d'élèves de lycées généraux et professionnels. Dans ce cadre, le CRÉDOC est évaluateur d'une expérimentation d'actions de sensibilisation à l'« esprit d'entreprendre » notamment par des exercices de mise en situation des élèves sur la création et la gestion d'entreprises virtuelles. Enfin, en avril 2010, 47 expérimentations ont été lancées pour aider les jeunes à obtenir le permis de conduire (appel à projets « 10 000 permis pour réussir »).

I.3. Les exigences méthodologiques de la Mission d'Animation du Fonds d'expérimentation pour la Jeunesse (MAFEJ)

Les dossiers soumis au Fonds d'expérimentation pour la Jeunesse devaient tenir compte des recommandations du guide méthodologique sur l'évaluation à l'intention des porteurs de projet. Ce guide de 22 pages a été rédigé en avril 2009, au moment du lancement de l'AP1. L'objectif de l'évaluation y est résumé par la formule suivante : « (…) Savoir si le programme améliore effectivement la situation des personnes qui en bénéficient » (p.1). Ce guide méthodologique prône le principe d'une évaluation externe et indépendante de l'expérimentateur. Les méthodes quantitatives, avec comparaison entre l'échantillon des bénéficiaires de la mesure et un échantillon témoin, sont présentées comme étant supérieures à toute autre méthode pour évaluer les effets du programme. La « randomisation » ou répartition aléatoire d'un échantillon de participants potentiels en un groupe témoin et un groupe bénéficiaire est considérée comme la méthode « la plus précise et la plus robuste » pour évaluer l'impact du programme (p.5), car elle permettrait de s'affranchir des biais de sélection sur les individus bénéficiant du programme (p.11). Ce n'est qu'en cas d'impossibilité d'adopter la randomisation que d'autres méthodes sont présentées comme envisageables, telles que la mise en place d'un échantillon témoin construit sur critères raisonnés.

Ce guide incite à la mise en place d'un système d'information par le porteur de projet pour suivre les populations concernées et les actions réalisées. Cet outil de type tableau de bord devait être pris en compte dans les analyses évaluatives. Les approches qualitatives sont par ailleurs envisagées pour identifier les systèmes de valeur des acteurs comme des bénéficiaires, rendre compte des attentes et des besoins, analyser les difficultés rencontrées et suivre le déroulement du processus de l'expérimentation.

Les jurys constitués par la MAFEJ étaient composés d'experts, d'acteurs de terrain et de scientifiques. A la suite d'une première phase d'analyse des dossiers reçus, ils ont souvent demandé des précisions pour s'assurer que l'évaluation permettrait de suivre la méthodologie des groupes de bénéficiaires et témoins par affectation aléatoire, considérée comme l'étalon de l'évaluation. Des demandes d'approfondissement ont également porté sur les objectifs des expérimentations, la définition de leur public cible et le contenu des actions. Les projets "retenus

avec modifications" ont fait l'objet d'une deuxième version retravaillée suite aux remarques émises par le jury (ex : baisse du budget, précision sur l'un ou l'autre aspect, amélioration du protocole)[29].

Encadré 1. Les critères de la MAFEJ pour analyser chaque projet d'expérimentation

1. Adéquation entre la réponse de l'expérimentation et la réponse à la problématique posée par l'appel à projets
2. Précision et cohérence de la méthodologie du projet et de son pilotage
3. Valeur ajoutée de l'expérimentation par rapport à l'activité habituelle de la structure
4. Planification « réaliste » de la durée de l'expérimentation et des actions
5. Viabilité de la structure porteuse du projet sur un plan administratif et financier
6. Qualité de l'ancrage territorial du projet, précision autour du périmètre d'intervention de l'expérimentation
7. Mobilisation de partenariats interinstitutionnels
8. Présence d'un évaluateur externe au projet
9. Précision de l'identification du ciblage de la population
10. Cohérence entre le coût du projet, le résultat attendu et le nombre de jeunes concernés
11. Mobilisation de cofinancements

Source : DJEPVA-MAFEJ - Appels à projets : mode d'emploi - http://www.experimentation.jeunes.gouv.fr/13-appel-a-projets-mode-d-emploi.html

La décision finale revenait à un conseil de gestion auquel participaient les administrations centrales intéressées (notamment emploi, éducation nationale, enseignement supérieur, budget, jeunesse, Outre-Mer...) et les partenaires privés impliqués dans la démarche[30]. A partir des conclusions des jurys, le conseil de gestion sélectionnait les projets retenus et décidait du montant des subventions attribuées par le fonds.

[29] Sur les 169 projets de l'AP1 finalement retenus, 99 ont été retenus avec modifications et 70 sans.

[30] La fondation TOTAL et l'Union des Industries et des Métiers de la Métallurgie (UIMM).

I.4. Les porteurs de projet : une implication forte des missions locales

Dans l'ensemble, 29% des projets financés par le FEJ entre 2009 et 2011 ont été portés par une association, et 27% par une mission locale[31]. Les missions locales occupent cependant une part plus importante dans l'échantillon des projets évalués par le CRÉDOC.

En effet, neuf des dix-huit projets évalués étaient portés par des missions locales, généralement par un regroupement de missions locales (tableau 1). Les projets coordonnés par le Conseil National des Missions Locales (CNML) ont été déployés sur plusieurs sites (sept dans un cas, huit dans l'autre) et impliquaient dans chaque site une missions locale généralement associée à d'autres partenaires. Deux projets étaient portés par un Conseil Général et mis en œuvre par des missions locales intervenant sur leur département. Les deux projets portant sur le microcrédit ont été mis en place par des Centres Communaux d'Action Sociale. Un projet concernant l'orientation des lycéens était porté par un rectorat. Par ailleurs, deux projets étaient portés par des structures associatives et un autre par une fondation. Enfin, une Chambre de Commerce et d'Industrie (CCI) était à l'origine d'un des projets.

Tous les projets concernant les jeunes sous main de justice associaient étroitement les services pénitentiaires d'insertion et de probation (SPIP).

Les cofinanceurs ont été des institutions variées : collectivités territoriales, associations, fonds propres du porteur de projet.

I.5. Les motivations des porteurs

En 2009, un premier Cahier de recherche du CRÉDOC sur les expérimentations sociales[32] distinguait parmi les raisons qui ont motivé les acteurs à répondre à l'appel à projets deux grandes rubriques :

- les motivations liées à un diagnostic local,
- les objectifs stratégiques et politiques à long terme.

Pour certains des projets ces deux types de motivation coexistent.

En tenant compte du contenu des projets et de leur différenciation par rapport à la situation antérieure sur le territoire, les projets peuvent être classés en deux grandes catégories :

- ceux qui visent à créer une innovation, soit en se concentrant sur un public qui n'était pas ou peu pris en compte par les structures impliquées dans l'action, soit en créant de

[31] Rapport d'activité 2009-2011 du Fonds d'Expérimentation pour la Jeunesse. Les missions locales sont des structures du service public de l'emploi accueillant des jeunes âgés de 16 à 25 ans et les accompagnant sur les questions d'emploi, de formation, de logement ou de santé

[32] Brézault, Olm, Angotti, Simon, 2009.

nouveaux modes d'intervention. Il est alors possible d'identifier un « service plus », une plus-value par rapport au droit commun ;

- et ceux qui consistent à intensifier des actions déjà existantes, soit en renforçant le rythme de suivi de l'accompagnement des publics jeunes, soit en cherchant à augmenter le nombre des bénéficiaires de ces actions.

Si, à la lecture des projets, certaines expérimentations semblent se classer dans la première catégorie, le déroulé du programme n'a pas toujours été à la hauteur des ambitions énoncées dans les projets. Faute de temps, de moyens humains ou suite à des difficultés non envisagées au stade de l'élaboration du programme pour développer des partenariats, le contenu des projets s'est parfois limité à un renforcement de l'accompagnement existant préalablement au financement par la MAFEJ.

I.6. Le contenu des projets

Les projets concernant les jeunes sous main de justice visaient à prévenir la récidive en favorisant l'insertion professionnelle des jeunes ayant connu l'incarcération. Ils s'appuyaient sur le développement ou le renforcement des partenariats entre les acteurs de l'insertion, notamment les missions locales et les services pénitentiaires. Ils comportaient des actions de formation, de séances de travail en groupe ou individuellement et certains projets comportaient un stage.

Deux des projets relevant de l'orientation s'adressaient à des jeunes scolarisés pour les éclairer dans leurs choix d'études, en les aidant à tenir compte à la fois de leurs aptitudes et motivations et en leur apportant une connaissance élargie des perspectives possibles (cursus et métiers). Deux autres projets s'adressaient à des décrocheurs scolaires ou universitaires et avaient pour objectif soit de leur permettre de réintégrer un cursus d'études (sous statut scolaire ou en alternance) soit de recevoir un soutien pour les aider à s'insérer professionnellement.

Les projets sur l'alternance visaient non seulement à orienter davantage de jeunes vers des contrats d'apprentissage ou de professionnalisation, afin d'améliorer leur niveau de qualification et leur offrir davantage de perspectives dans leur carrière professionnelle, mais aussi à sécuriser les parcours en alternance en limitant les risques d'abandon par un suivi après la signature du contrat d'alternance. Les bénéficiaires ciblés par ces expérimentations étaient souvent des jeunes qui avaient abandonné les études et cherchaient à s'insérer dans la vie professionnelle, alors que la grande majorité des apprentis proviennent généralement directement du système scolaire.

Les expérimentations dans le domaine du logement avaient pour objectif de favoriser l'accès au logement autonome des jeunes en insertion professionnelle. Elles reposaient sur l'hypothèse qu'un accompagnement renforcé permet d'améliorer l'accès au logement des jeunes, et de favoriser leur insertion sociale et professionnelle. Au-delà de l'information et du suivi de ces jeunes, la dimension partenariale était forte dans ces projets ; elle permettait de mobiliser davantage de logements et permettre l'accès des publics jeunes à un plus grand nombre de solutions dans le domaine de l'habitat.

Les deux projets sur l'accès au microcrédit personnel cherchaient à soutenir les jeunes dans leurs projets d'insertion socioprofessionnelle et à leur apporter un accompagnement par des professionnels de l'économie sociale et familiale pour éviter tout risque de fragilisation en cas de difficultés budgétaires.

I.7. Une forte sollicitation du CRÉDOC pour évaluer des projets

Le CRÉDOC a été sollicité par de nombreux porteurs de projet pour évaluer en tout 45 projets.

Pour 22 de ces projets, le CRÉDOC ne s'est pas porté évaluateur, soit parce que le porteur de projet n'avait pas trouvé de partenaire pour cofinancer l'expérimentation avant la date limite de l'envoi, soit parce que le CRÉDOC a considéré que le projet ne remplissait pas les conditions suffisantes pour être évaluable (objectifs peu lisibles, différence insuffisamment marquée avec les programmes déjà existants sur le territoire), soit encore parce que les délais étaient trop courts pour pouvoir élaborer un protocole d'évaluation satisfaisant.

Il est probable que nombre de projets n'ont pu être présentés auprès des jurys de la MAFEJ faute d'avoir pu trouver un évaluateur pour monter un protocole d'évaluation dans le temps imparti. Les évaluateurs se sont trouvés en position de faire une première sélection des projets avant l'intervention des jurys.

Finalement, le CRÉDOC a été associé en tant qu'évaluateur à 23 projets d'expérimentation qui ont été soumis au jury. Parmi eux, 16 ont été acceptés.

I.8. Les propositions méthodologiques du CRÉDOC pour l'évaluation : une mixité de méthodes

Le CRÉDOC a généralement proposé des méthodes mixtes, combinant des étapes quantitatives et des campagnes d'entretiens qualitatifs pour analyser l'impact des projets sur le devenir des jeunes. Ces méthodes ont été adaptées à chaque projet et répondaient aux attentes de la Mission d'Animation du Fonds d'expérimentation pour la Jeunesse exprimées dans le guide méthodologique. Ainsi, dans la grande majorité des projets auxquels le CRÉDOC a été associé, l'évaluation reposait sur :

- deux ou trois interrogations par questionnaire auprès des bénéficiaires et d'un groupe témoin, portant sur les parcours antérieurs, sur les actions suivies, sur leur devenir pendant et après le temps de l'expérimentation. Pour les bénéficiaires, les questionnaires portaient également sur leurs attentes et leur satisfaction par rapport au programme ;
- un tableau de bord rempli par les professionnels suivant les jeunes, afin de mesurer aussi exhaustivement que possible les actions suivies, les propositions faites, le rythme réel de l'accompagnement ;

- deux campagnes d'entretiens qualitatifs auprès d'acteurs (une en début de programme et une autre en fin de programme) pour recueillir le point de vue des porteurs de projet et des partenaires et analyser l'évolution du partenariat pendant la durée de l'expérimentation.

L'analyse des questionnaires devait permettre une vision assez complète du parcours des jeunes avant leur entrée dans le dispositif, pendant l'expérimentation et révéler le regard des jeunes sur les actions suivies et la pertinence du dispositif. Le tableau de bord est un outil a priori exhaustif pour mesurer l'ensemble des actions effectivement proposées et suivies dans le cadre de l'expérimentation, et l'intensité de l'accompagnement. Les entretiens qualitatifs des acteurs se prêtent bien à faire émerger les différences de point de vue et d'attentes entre les partenaires, les difficultés rencontrées, et l'évolution du partenariat.

Au moment des réponses à l'appel à projets, une grande majorité des évaluations envisageaient de faire appel à la randomisation pour distinguer groupe bénéficiaire et groupe témoin. Les premiers mois de mise en place des projets feront apparaître des difficultés dans la mise en œuvre de cette méthode, ce point étant développé dans le chapitre suivant.

Encadré 2. Les critères de la MAFEJ pour analyser chaque protocole d'évaluation

1. Clarté et pertinence des objectifs de l'évaluation par rapport aux buts poursuivis par le projet
2. Précision dans la présentation du protocole d'évaluation
3. Précision dans le mode de suivi des bénéficiaires (données utilisées, enquêtes ou entretiens prévus)
4. Précision des indicateurs quantitatifs pressentis par rapport au degré d'atteinte des objectifs poursuivis
5. Précision des indicateurs qualitatifs pressentis par rapport au degré d'atteinte des objectifs poursuivis
6. Mise en place d'une comparaison d'un groupe de bénéficiaires et d'un groupe témoin
7. Rigueur du mode de détermination du groupe témoin
8. Taille des échantillons mobilisés et leur précision statistique pour conclure quant à l'efficacité du dispositif évalué
9. Connaissance du contexte scientifique dans lequel s'inscrit l'évaluation
10. Le coût de l'évaluation
11. Capacité du protocole à fournir des résultats conclusifs quant à l'efficacité et au caractère généralisable des dispositifs

Source : DJEPVA-MAFEJ - Appels à projets : mode d'emploi - http://www.experimentation.jeunes.gouv.fr/13-appel-a-projets-mode-d-emploi.html

I.9. Un temps très court pour rédiger les protocoles d'évaluation

Le délai de moins de trois mois qui a séparé l'appel à projets de la date de remise des propositions s'est souvent révélé très court, pour les porteurs de projet comme pour les évaluateurs. Certains porteurs de projet n'avaient alors pas une idée très précise du public qui serait touché par l'expérimentation, ni en termes de volume, ni en termes de profil. Les actions à mettre en place n'étaient pas toujours très précisément définies. Au moment où l'opportunité d'élaborer un protocole d'expérimentation s'est créée, il leur a fallu un certain temps pour consulter des partenaires, trouver des cofinanceurs et développer des idées d'actions innovantes à destination des jeunes. En conséquence, les sollicitations des porteurs de projet pour associer le CRÉDOC à l'évaluation de leur proposition d'expérimentation ont souvent eu lieu peu de temps avant la date finale d'envoi des dossiers, et ceci même quand les porteurs avaient une idée déjà bien élaborée du contenu de l'expérimentation. Il serait souhaitable d'associer l'évaluateur au projet le plus en amont possible, afin de s'assurer que l'expérimentation revêt un contenu susceptible de donner lieu à une évaluation rigoureuse.

II. LE LIEN ENTRE PORTEUR ET EVALUATEUR AU STADE DE LA REDACTION DU PROJET

II.1. L'évaluateur contribue à préciser les objectifs du projet

Le CRÉDOC en tant qu'évaluateur n'a généralement pas été associé au montage de projet, ayant été sollicité quand le projet était déjà construit. Il a suivi le principe promu par le guide d'évaluation du FEJ d'une posture externe de l'évaluateur.

Toutefois, pour être en capacité de mesurer les effets de l'expérimentation sur le devenir des jeunes bénéficiant de l'action, le CRÉDOC a parfois dû pousser le porteur à remodeler son projet. Dans ces cas plusieurs allers-retours entre porteur et évaluateur ont été nécessaires pour que le projet soit suffisamment précis, sa cible identifiée et les actions prévues bien développées et différenciées du droit commun existant déjà sur le territoire.

L'évaluateur se retrouve dans une situation ambiguë et parfois contradictoire : il cherche à garder son indépendance face au porteur de projet et à ne pas intervenir sur le contenu du projet pour en respecter la logique interne, tout en étant amené à formuler des préconisations pour que le projet soit évaluable. Le porteur de projet, quand son action est encore imprécise, est parfois amené à consulter l'évaluateur sur la conduite même du projet, sur ses objectifs. Dans ces situations, le CRÉDOC a cherché à pousser le porteur à préciser son projet, tout en évitant de se substituer à lui dans les choix opérés.

Certains consultants positionnés sur le champ de l'évaluation remplissent également des fonctions d'aide à l'élaboration de projet. La combinaison des deux rôles est de nature à mettre en question l'extériorité de l'évaluation. Quand une aide au montage de projet est nécessaire, il serait souhaitable que deux structures différentes interviennent, l'une en soutien du porteur pour la conception du projet, l'autre pour l'étape d'évaluation. C'est ainsi qu'on peut analyser le projet d'expérimentation concernant des décrocheurs universitaires dans l'agglomération de Dunkerque, « De l'université à l'entreprise ». Le porteur de projet est l'association « Entreprendre Ensemble », dont l'une des fonctions est le pilotage de la mission locale. Ce projet a été évalué conjointement par le CRÉDOC et un bureau d'études du Nord, E2I. Ce bureau d'études intervient régulièrement en soutien au montage de projet auprès d'Entreprendre Ensemble. E2I a conduit son évaluation sur l'université, alors que le CRÉDOC se chargeait des entretiens auprès des salariés de la mission locale. Ainsi, l'extériorité de l'évaluateur par rapport au porteur est conservée.

Il peut arriver que l'évaluateur participe à la conception même du projet. Dans le cas d'un projet concernant la mobilisation de décrocheurs scolaires « en errance », le rôle de conseil de l'évaluateur était explicitement prévu, comme le montrent les extraits suivants de la convention tripartite entre le Ministère, l'évaluateur (le CRÉDOC) et le porteur de projet : « L'évaluation sera conduite sur l'ensemble de la période avec une étape préalable de construction des outils de suivi et d'évaluation en concertation avec les partenaires du projet. Cette étape préalable comprend la définition, avec l'ensemble des partenaires, de ce qui signifie « être en errance », la définition d'indicateurs pour appréhender ceux qui sont déjà en situation d'errance comme ceux qui sont en

passe de l'être, et enfin la définition des objectifs à atteindre pour les missions locales impliquées. »

Dans le cas d'un des programmes concernant des jeunes sous main de justice s'adressant à des détenus de 18 à 25 ans où six groupes de jeunes entraient tous les deux mois dans l'action, le CRÉDOC est intervenu auprès du porteur pour augmenter les effectifs de jeunes concernés afin que l'évaluation puisse être établie sur au moins 70 bénéficiaires. Ce seuil semblait un minimum pour qu'une analyse quantitative des questionnaires puisse être menée.

De par sa connaissance des travaux sociologiques et économiques sur le public visé par une expérimentation, l'évaluateur peut apporter un éclairage sur l'état de la question au porteur de projet, plus près du terrain mais plus éloigné des critères pertinents d'analyse. Un projet d'accompagnement de travailleurs pauvres s'est appuyé sur l'expertise du CRÉDOC pour définir plus précisément la cible de son intervention : quels étaient les critères pour considérer qu'une personne était « travailleur pauvre », quels étaient les différents profils de travailleurs pauvres que l'expérimentation cherchait à accompagner ?

II.2. L'intervention des porteurs de projet sur les méthodes évaluatives : des raisons éthiques et pratiques

Le porteur peut également avoir une influence sur les méthodes d'évaluation, par exemple en se montrant réticent à certaines méthodes d'évaluation – telles que la randomisation ou la mise en place d'un tableau de bord *ad hoc*. Des raisons éthiques peuvent être à l'origine de ces réticences, en particulier quand la sélection aléatoire du public amène à priver du bénéfice de l'expérimentation une partie d'un public potentiel. Les raisons sont parfois pratiques, quand les exigences de l'évaluateur en matière de constitution d'outils de suivi entraineraient une mobilisation forte des professionnels accompagnant les jeunes dans le projet, et qu'il semble au porteur de projet impossible de faire accepter dans les délais impartis cette charge supplémentaire de travail à l'équipe de professionnels intervenant sur l'action. Il s'ensuit des étapes de négociations où les critères scientifiques ne sont pas toujours les plus déterminants dans les choix opérés in fine.

Les porteurs de projet ont toujours été consultés par le CRÉDOC dans l'élaboration des outils de suivi de l'évaluation (questionnaires auprès des bénéficiaires, listes d'acteurs à interviewer, mise au point de tableaux de bord). L'association des acteurs en interaction avec les bénéficiaires de l'action ainsi qu'avec les partenaires permet de compléter la liste des actions possibles et d'affiner le regard de l'évaluateur sur les éléments pertinents à analyser dans le parcours des jeunes avant, pendant et après l'expérimentation. Les acteurs savent également quelles sont les informations disponibles à intégrer dans un tableau de bord, et celles qui seront difficiles à renseigner. Cette participation des acteurs de terrain aux outils de l'évaluation n'est pas limitée aux expérimentations. Dans ce cas on peut parler de co-construction des outils de l'évaluation. Cette

étape était explicitement prévue dans le cas de la convention déjà évoquée concernant les décrocheurs.

Les échanges avec le CRÉDOC sur les outils méthodologiques d'évaluation ont souvent permis la création d'instruments de suivi de l'action mobilisés par les porteurs eux-mêmes. Les porteurs de projet ont presque toujours réalisé une évaluation interne pour suivre les actions mises en place et ont pu mobiliser ces outils dans ce cadre.

II.3. Associer un facilitateur

En cas de nécessité, les porteurs peuvent s'appuyer sur un tiers pour le montage de leur projet ou pour le pilotage. Un tiers, intermédiaire entre le porteur et les partenaires, peut faciliter le déroulement du projet, la centralisation des informations voire l'organisation concrètes de rencontres. Cela peut s'avérer utile quand le développement de partenariats entre acteurs qui se connaissent peu fait partie des objectifs. Ce tiers apporte une plus-value sensible à la mise en œuvre des projets quand ceux-ci comportent une dimension multi-sites, notamment pour apporter une plus grande homogénéité au contenu des actions.

De par son expérience dans l'accompagnement de la conduite de projet dans le champ des politiques sociales, l'Agence Nouvelle des Solidarités Actives (ANSA) a joué ce rôle dans les deux projets portés par le CNML que le CRÉDOC a évalués.

III. LA MISE EN ŒUVRE DES EXPERIMENTATIONS : DIFFICULTÉS RENCONTRÉES ET RECOMMANDATIONS

La mise en œuvre des expérimentations a donné lieu à diverses modifications des projets et à de nombreuses adaptations des dispositifs par rapport à leur forme initiale. Le guide méthodologique de la MAFEJ à destination des porteurs n'aborde que brièvement ces aspects. Or, des difficultés très diverses ont pu surgir en cours d'expérimentation. Cette partie décrit les principales difficultés qui ont été rencontrées au début et en cours d'expérimentation et illustrent la complexité de l'exercice d'expérimentation. Nous mettons ici en exergue les obstacles pratiques qui se sont présentés. Un certain nombre d'entre eux peuvent néanmoins être évités dès lors que porteurs et évaluateurs partent avertis et les anticipent.

III.1. Une mobilisation du public parfois plus lente que prévue : la question de la montée en charge des dispositifs

La plupart des projets présentaient des objectifs chiffrés de jeunes à intégrer dans l'expérimentation. Dans l'ensemble une faible part seulement des expérimentations se sont déroulées conformément aux prévisions en termes de volume, avec un nombre de bénéficiaires correspondant aux prévisions dès le démarrage de l'évaluation.

Le problème rencontré dans plusieurs cas est celui de la montée en charge des projets. Certains projets visant un public restreint, comme ceux concernant les jeunes sous main de justice, prévoyaient cependant dès le départ de faire entrer des jeunes dans le dispositif soit progressivement, au fil de l'eau, soit successivement par groupe de dix par exemple, et comportaient un protocole d'évaluation adapté au rythme d'entrée.

Dans les projets qui visaient des **publics peu connus des structures**, la mobilisation a parfois été nettement plus difficile que prévu. C'est ce qui s'est produit dans le projet d'accompagnement des travailleurs pauvres qui visait à les aider à se stabiliser sur le plan professionnel en vue d'atteindre un niveau de revenu leur permettant de sortir de la pauvreté. C'est un public qui n'est que rarement en contact avec les structures d'aide à l'insertion, qui est déjà dans la vie professionnelle avec des horaires atypiques et qui a appris à trouver des solutions, certes précaires, par ses propres réseaux. L'importance du non-recours au Revenu de Solidarité Active (RSA) dans son volet activité montre bien la difficulté à toucher cette population de travailleurs pauvres. C'est ce qui explique qu'une partie de l'énergie des chargés de projet a été consacrée principalement au recrutement et non à l'accompagnement. Un autre projet qui cherchait à inciter des jeunes déscolarisés et sans insertion professionnelle à intégrer des actions de formation en alternance a rencontré également des difficultés pour mobiliser le public potentiel, les jeunes se

présentant aux réunions collectives ne manifestant pas toujours leur intérêt pour participer au programme expérimental.

C'est parfois la **difficulté à impliquer un partenaire** qui n'était pas associé au projet dès le début de l'expérimentation et sur qui le porteur comptait pour identifier son public qui est à l'origine d'une faible montée en charge. Ces situations se sont produites dans un des projets visant des décrocheurs scolaires éloignés des institutions d'aide à l'insertion. Il était prévu que pendant les premiers mois du projet le Rectorat fournisse des listes de jeunes sortis sans diplôme – ou au plus avec un brevet des collèges – et non réinscrits dans un établissement scolaire. Les délais pour établir un partenariat efficace ont été tels que le projet a surtout intégré des jeunes inscrits dans le passé à la mission locale mais n'ayant plus aucun contact depuis plusieurs mois. D'autres sources ont pu être mobilisées grâce au partenariat avec des clubs de prévention intervenant sur le territoire des missions locales. Ces difficultés se sont cependant traduites par un nombre de personnes bénéficiaires des actions inférieur à celui prévu au stade du projet.

Autre cas de figure : lorsqu'un **partenariat a mis du temps ou a peiné à se mettre en place** comme pour l'accompagnement d'étudiants de l'université à l'entreprise. La mission locale devait obtenir des listes de jeunes étudiants absentéistes grâce à la mise en place d'un repérage au niveau de l'université. Mais la coordination n'a pas été bonne et l'université n'ayant pas mis en place un dépistage précoce des étudiants absentéistes, les entrées dans l'expérimentation se sont faites au compte-gouttes, au rythme de l'arrivée des jeunes à la mission locale par d'autres voies que l'université.

La contrainte de faire entrer un nombre donné de bénéficiaires dans le projet, qui plus est dans un temps limité, a ainsi parfois conduit à un **élargissement du public cible** : jeunes en milieu ouvert dans une expérimentation en faveur des jeunes sous main de justice prévue au départ pour des jeunes incarcérés, mobilisation de jeunes déjà inscrits dans une mission locale quand le projet initial visait à s'adresser à des jeunes qui ne connaissaient pas cette structure, ou encore intégration dans le programme d'accompagnement vers le logement de tous les jeunes se présentant à la Mission locale, même si leur besoin en logement n'était pas urgent.

De façon plus générale, une **durée** de deux-trois ans apparaît **trop courte** pour procéder à la fois à la création du projet, sa finalisation et son lancement, sa montée en charge et l'atteinte du rythme de croisière[33]. C'est particulièrement le cas quand les effectifs des bénéficiaires ne sont pas massifs, soit parce que l'on étudie un territoire restreint, soit parce que le phénomène étudié est rare ou discontinu, ou encore parce que les personnes éligibles sont difficiles à repérer (par exemple, les décrocheurs en rupture d'accompagnement, qui, par définition, sont difficiles à identifier[34]). Dans ces cas-là, il nous semble indispensable de prévoir une période pilote de montée en charge du projet avant le démarrage de l'évaluation quantitative. Une telle phase peut en revanche faire l'objet d'évaluation qualitative afin d'analyser les conditions de mise en œuvre. Dans

---

[33] Dans l'un des derniers appels à projets d'expérimentations de la Commission Européenne, la durée consacrée à l'expérimentation elle-même est d'ailleurs passée de deux à trois ans, illustration de la prise de conscience de ce point.

[34] On parle d'« errance institutionnelle » pour qualifier ces jeunes qui décrochent des dispositifs destinés à les « raccrocher ».

les expérimentations pour la jeunesse, la phase de montée en charge et les premières interrogations n'avaient généralement pas été dissociées.

III.2. Le développement des partenariats et des actions expérimentées plus long que prévu

L'intégralité des actions prévues lors des réponses à l'appel d'offres n'ont pas toujours pu être menées, par manque de temps de la part des porteurs de projet ou encore par l'insuffisante implication de certains partenaires. Un exemple typique de ces difficultés à monter des actions prévues est celui du programme concernant des décrocheurs universitaires cités ci-dessus. Le nombre de jeunes éligibles est moins élevé que prévu car ils n'ont pas été identifiés au niveau de l'université. En conséquence il n'a pas été possible de créer les ateliers spécifiques prévus, car le nombre de jeunes de niveau d'études supérieur au baccalauréat était insuffisant.

Dans un projet sur l'alternance, les entrées directes ont été moins nombreuses que prévues car certains centres de formation ont exprimé des réticences pour faire accompagner des jeunes déjà inscrits en formation dans leur structure par un partenaire extérieur pour un suivi renforcé alors qu'ils proposaient déjà, en interne, un accompagnement pour prévenir les ruptures.

Par ailleurs, un des programmes concernant les jeunes sous main de justice prévoyait de développer un partenariat fort avec des entreprises. Faute de temps et de moyens, celui-ci n'a cependant pas pu être mis en place dans la durée de l'expérimentation. Autre exemple : un projet sur le thème de l'alternance prévoyait une phase d'accompagnement dans l'emploi en vue de sécuriser le parcours de jeunes après l'obtention de leur diplôme ; cependant très peu de jeunes ont atteint cette phase dans la durée impartie à l'expérimentation, donc l'accompagnement a très peu eu lieu à ce stade du parcours.

III.3. Des recrutements temporaires parfois difficiles à effectuer

Recruter des salariés formés aux métiers spécifiques s'avère parfois difficile. Une plateforme d'aide à l'accès au logement souhaitait ainsi recruter en CDD des conseillers en économie sociale et familiale pour le temps de l'expérimentation. Or, localement, ces profils étaient en position favorable sur un marché de l'emploi où la demande dans ces postes était supérieure à l'offre, et il a été long pour les porteurs de mettre en place l'équipe dédiée au projet.

III.4. Pour des études de faisabilité préalables et bien cadrées

L'essentiel des réajustements tant dans l'expérimentation que dans la conduite de l'évaluation a eu lieu après le démarrage du programme, les difficultés apparaissant à ce stade. D'où l'intérêt de procéder à une étape d'étude de faisabilité[35] pour avoir une vision plus claire des rythmes de montée en charge des dispositifs et des difficultés éventuelles à aplanir pour le développement de l'action. A noter, les études de faisabilité sont à mener sur les territoires concernés par l'expérimentation.

Trois projets ont effectivement été précédés d'une étude de faisabilité. Deux de ces expérimentations incluent parmi les porteurs le Conseil National des Missions Locales (CNML), l'une sur l'apprentissage et l'autre sur l'accès au logement. La troisième expérimentation concernait le développement de l'alternance. Les deux projets portés par le CNML concernaient un nombre important de sites. L'étude de faisabilité a permis d'analyser les conditions de mise en œuvre de l'expérimentation sur les différents sites. Les discussions entre les porteurs des projets (des grandes fédérations de structures locales accompagnant les jeunes), les responsables des missions locales chargées de l'étape opérationnelle auprès des jeunes et les évaluateurs ont été l'occasion :

- de convaincre les partenaires de l'utilité de procéder à un tirage au sort pour séparer les bénéficiaires des témoins,
- de s'accorder sur les modes de repérage des jeunes, sur la forme prise par les plateformes et sur les objectifs de chaque expérimentation.

L'étape de faisabilité de l'expérimentation sur l'alternance a connu pour limite son calendrier restreint. Cette expérimentation portait sur trois sites dans des régions différentes, et au moment de l'étude de faisabilité une seule région avait commencé à nouer un réseau de partenaires. L'étude de faisabilité n'a donc pas permis de mesurer l'adhésion des partenaires potentiels au projet. Elle a été l'occasion de préciser les modalités de sélection des bénéficiaires et des témoins. Les méthodes mises au point lors de cette phase de faisabilité n'ont toutefois pas pu être maintenues à cause de la forte déperdition des jeunes entre l'étape des réunions d'information et celle de l'entrée effective dans le programme.

Les étapes de faisabilité devraient pouvoir vérifier que les projets prévus correspondent à un réel besoin pour le public visé, et permettre d'estimer les effectifs potentiels de bénéficiaires.

---

[35] Une étape déjà préconisée par le Cahier de Recherche du CRÉDOC de décembre 2009 qui tirait un premier bilan des expérimentations sociales en direction des jeunes, quelques mois après leur mise en place (Brézault, Olm, Angotti, Simon, 2009).

III.5. Des relations de confiance indispensables entre porteur de projet et évaluateur

La mise en œuvre de l'évaluation repose fortement sur la transmission d'informations par le porteur du projet en vue de connaitre :

- les actions effectivement mises en place,
- les coordonnées des bénéficiaires du programme et du groupe témoin,
- les partenaires impliqués.

L'une des évaluations démarrées par le CRÉDOC n'a pu se poursuivre par manque de coopération du porteur de projet dans le processus d'évaluation. Le CRÉDOC n'a pas pu disposer des informations utiles pour élaborer un diagnostic concernant l'impact du projet sur le parcours des bénéficiaires. Le porteur de projet a exprimé des doutes sur la capacité de l'évaluateur de mener des investigations sans fragiliser le partenariat avec les principaux acteurs et la relation avec les jeunes accompagnés. Le porteur de projet percevait l'évaluation comme une contrainte, voire comme un frein à la mise en œuvre de son projet. Le principe même d'une évaluation externe n'a pas été accepté par le porteur. Ce climat de défiance a empêché la transmission d'informations essentielles : les coordonnées des partenaires de terrain, le tableau de bord répertoriant les jeunes entrés dans le dispositif. Un très faible nombre des questionnaires conçus par le CRÉDOC et distribués par le porteur ont été communiqués. Malgré l'intervention de la MAFEJ, et plusieurs réunions pour expliquer les objectifs et méthodes de l'évaluation, et malgré des propositions d'ajustements méthodologiques du protocole évaluatif, la mission d'évaluation a dû être interrompue.

PARTIE 2 : REGARD CRITIQUE SUR LES METHODES D'EVALUATION MISES EN ŒUVRE

Treize des projets évalués par le CRÉDOC présentaient un dispositif d'évaluation construit suivant l'ensemble des recommandations du guide méthodologique de la MAFEJ et incluaient :

- un échantillon bénéficiaire et un échantillon témoin (constitués aléatoirement dans 10 des projets et constitué de façon raisonnée dans 3 projets) ;
- deux à trois interrogations quantitatives auprès des jeunes bénéficiaires et témoins ;
- deux à trois interrogations qualitatives auprès des partenaires ;
- la mise en place d'un tableau de bord de suivi de la situation des jeunes et des actions réalisées.

Dans les autres projets, seule une évaluation qualitative auprès des acteurs et des jeunes était programmée. En particulier, pour les projets portant sur le microcrédit, les effectifs de bénéficiaires prévus étaient très faibles et la constitution d'un échantillon de témoins n'ayant pas recours au micro-crédit difficilement réalisable ont conduit à écarter d'emblée les approches quantitatives.

En pratique, la mise en place d'un échantillon témoin (aléatoire ou non) et la diffusion des questionnaires auprès des jeunes a présenté des difficultés, en grande partie liées à la très petite taille des expérimentations. Cette partie décrit les difficultés rencontrées dans la mise en œuvre des méthodes quantitatives et les alternatives déployées. En particulier, nous rappelons que les méthodes qualitatives ont toute leur place dans l'évaluation des expérimentations sociales. Les entretiens (ainsi que les tableaux de bord collectés) ont en effet permis de compenser, en partie, les limites des enquêtes sur très petits échantillons.

I. DES APPROCHES QUANTITATIVES AMBITIEUSES A ADAPTER A LA TAILLE DES EXPÉRIMENTATIONS

Le guide méthodologique pour l'évaluation des expérimentations sociales du Fonds d'expérimentation pour la Jeunesse présente la méthode d'expérimentation contrôlée et randomisée comme étant supérieure aux autres méthodes d'évaluation quantitative. Cette méthode consiste en la mise en place par tirage aléatoire de deux groupes : un groupe bénéficiaire du dispositif innovant et un groupe témoin (non bénéficiaire du dispositif). Ce tirage permet d'obtenir des échantillons sans biais de sélection et donc en théorie un échantillon témoin ayant les mêmes caractéristiques observables et inobservables que l'échantillon bénéficiaire. L'efficacité du dispositif est alors évaluée en comparant directement l'évolution des situations de deux échantillons à l'issue de l'expérimentation.

I.1. Groupes bénéficiaire et témoin : l'idéal de la randomisation

I.1.1. Un tirage aléatoire en pratique souvent difficile à mettre en œuvre

Le refus du tirage aléatoire par les porteurs de projet pour raison éthique

Pour des acteurs du champ social, il est injuste de ne pas offrir le même accompagnement à deux jeunes ayant les mêmes besoins et de ne pas proposer le dispositif expérimental à l'ensemble de la population cible qui par définition a besoin de ce dispositif. La majorité des porteurs de projets se sont donc, dans un premier temps, montrés réticents envers cette méthodologie. Le premier frein à la mise en place du tirage aléatoire a donc été de nature éthique. Deux principaux arguments sont généralement avancés afin de répondre à cette critique :

- L'évaluation contrôlée est la méthode scientifique la plus simple et la plus efficace pour observer les effets du dispositif expérimental par rapport au droit commun. Elle permet de mettre en avant de manière indiscutable les impacts du dispositif sur la situation des bénéficiaires. De plus, les témoins continuent de bénéficier du droit commun, l'évaluation ne conduit pas à abandonner leur suivi habituel.
- Les moyens financiers, techniques ou humains dédiés à l'expérimentation pouvant être limités, il s'avère impossible d'intégrer tous les jeunes du public cible dans le dispositif. Dans ce cas, le tirage aléatoire devient un moyen « juste » de sélectionner les bénéficiaires du dispositif. Les porteurs de projet n'ont plus à porter la responsabilité du choix du public qui bénéficiera du dispositif.

Malgré ces arguments, deux porteurs de projet ont refusé le tirage aléatoire des échantillons bénéficiaire et témoin. Dans ces deux cas, les échantillons témoins ont été constitués sur des territoires géographiquement proches de ceux concernés par l'expérimentation et dont les caractéristiques étaient a priori similaires[36].

[36] L'évaluation du RSA expérimental, dispositif souvent présenté comme le premier exemple à grande échelle d'expérimentation contrôlée en France, n'a finalement pas appliqué la méthode du tirage au sort des bénéficiaires. Les dix départements volontaires pour mener l'expérimentation n'étaient pas favorables à cette méthode, et ont proposé une répartition des zones entre territoires expérimentaux et territoires témoins. Cependant, les critères de répartition des territoires entre l'une ou l'autre de ces catégories n'ont pas été clairement explicités, et n'étaient pas homogènes d'un département à l'autre (Gomel, Séverin, 2011).

Des effectifs parfois trop faibles pour obtenir des résultats statistiquement significatifs

Lors de la réponse à l'appel à projets, les porteurs et l'évaluateur se sont fixés un effectif cible. Le public devait alors être suffisamment important pour permettre la mise en place d'un tirage aléatoire et la création d'un échantillon bénéficiaire et d'un échantillon témoin. Il s'est avéré que certains projets avaient anticipé un effectif potentiel de bénéficiaires supérieur à celui réellement mobilisable. Cela s'est produit dans le cas d'un projet dont le but était de favoriser l'insertion des jeunes à leur sortie de maison d'arrêt. Une fois pris en compte les critères de recevabilité d'une demande d'aménagement de peine par les juges (absences d'autres affaires en cours, disponibilité de places en semi-liberté, pas de problèmes comportementaux en détention…) et une fois écartés les jeunes en attente de jugement, les effectifs étaient très faibles. En effet, les jeunes susceptibles d'intégrer l'expérimentation n'étaient pas en nombre suffisant pour construire un échantillon de bénéficiaires utilisant l'ensemble des places offertes à l'intégration du dispositif et un échantillon témoin de même taille. La mise en place d'une évaluation contrôlée avec groupe témoin obtenu par randomisation n'était plus possible voire mettait en danger l'expérimentation (trop peu de bénéficiaires).

Dans d'autres cas, le principe même du tirage au sort d'un échantillon témoin a été rendu caduc du fait de la montée en charge nettement plus lente que prévue du dispositif. Sur un projet qui visait à promouvoir l'apprentissage auprès des jeunes éloignés de l'emploi, les échantillons bénéficiaires et témoins devaient être sélectionnés aléatoirement, selon la procédure suivante : les prescripteurs du dispositif orientaient les jeunes repérés comme potentiellement bénéficiaires et intéressés par le dispositif vers la plate-forme. Les jeunes étaient ensuite invités à une réunion d'information collective. Le tirage aléatoire était mis en œuvre parmi les jeunes exprimant leur motivation à l'issue de la réunion d'information collective. Néanmoins, la lenteur de la montée en charge du dispositif a freiné la mise en œuvre du protocole d'évaluation. Les orientations par les prescripteurs potentiels ont d'abord été très limitées (une trentaine de jeunes présents aux premières réunions d'information collective, là où environ 60 jeunes étaient attendus). Après les premières sélections aléatoires réalisées, de 15 à 20 jeunes étaient orientés vers le dispositif. Beaucoup ne l'ont finalement pas intégré. Au final, l'effectif était trop faible pour permettre la mise en œuvre des actions collectives telles qu'elles étaient initialement prévues. Il a donc été décidé dans un premier temps de retarder le début de l'évaluation afin de laisser le temps au dispositif de se mettre en place dans de bonnes conditions. Le dispositif connaît finalement, au moins en partie, le volume attendu: le nombre de jeunes participants aux séances d'information collectives est plus élevé (plus de 40 jeunes). Cependant, parmi ces jeunes, y compris lorsqu'ils expriment une motivation initiale, moins d'un tiers intègre effectivement le dispositif. Cette intégration a lieu après un effort très important des conseillers pour mobiliser les jeunes.

La mise en place d'un échantillon témoin par sélection aléatoire n'apparaît alors plus possible car :

- Si la sélection a lieu avant l'effort de mobilisation des jeunes : il n'est pas possible de considérer comme bénéficiaires tous les jeunes s'étant vus proposer le dispositif, la proportion d'un tiers de jeunes l'intégrant réellement et le faible nombre de bénéficiaires, ne permettant pas ensuite de corriger cette non-participation. Si l'évaluation ne conserve que les jeunes effectivement bénéficiaires, l'échantillon témoin est biaisé par le fait que les deux tiers d'entre eux n'auraient pas intégré le dispositif s'il leur avait été proposé.
- Il n'est pas possible d'opérer la sélection après effort de mobilisation des jeunes, cette mobilisation étant très coûteuse en temps pour les conseillers. Il n'est pas envisageable que dans la moitié des cas, elle aboutisse ensuite à un refus fait aux jeunes d'intégrer le dispositif.

Au final, les délais plus importants que prévus dans la phase de montée en charge ont amené à une réduction du nombre de jeunes effectivement suivis par le programme (un peu moins de 400 contre plus de 530 envisagés lors de la réponse à l'appel d'offres) et ont finalement conduit à l'abandon du tirage aléatoire et de l'échantillon témoin.

Un risque de contagion entre les deux échantillons si le dispositif n'est pas « étanche »

Même dans le cas où les échantillons ont été constitués aléatoirement, il s'est avéré impossible de relever des résultats différenciés entre bénéficiaires et témoins. La « contagion » du dispositif d'un échantillon à l'autre existe, par exemple, quand les témoins ont le même conseiller que les bénéficiaires et ont pu accéder au même accompagnement que ces derniers. Ce phénomène a été à l'origine de l'impossibilité de conclure de l'efficacité d'un dispositif. Ce problème a été rencontré notamment sur la thématique du logement : quand les conseillers avaient une opportunité de logement à proposer à un bénéficiaire mais qu'aucun ne se montrait intéressé, le logement était parfois proposé à un jeune du groupe témoin. Il était alors impossible d'obtenir des effets différenciés selon les échantillons, les deux publics ayant eu, au moins en partie, le même accompagnement.

La mise en place d'un tirage aléatoire des groupes bénéficiaires et témoins a donc souvent été difficile voire impossible et ce pour plusieurs raisons : la faiblesse des effectifs et le risque de contagion entre les échantillons bénéficiaire et témoin quand les mêmes acteurs accompagnent les deux groupes et que le dispositif couvre un territoire entier.

D'autres solutions ont alors été envisagées et mises en place mais elles aussi présentent des limites.

Encadré 3. Taille d'échantillon et puissance statistique

Au moment d'élaborer une méthodologie d'enquête quantitative, une question revient en permanence, celle de la **taille minimale que doit avoir l'échantillon** de personnes interrogées pour obtenir des résultats « suffisamment précis ». Ce « suffisamment précis » dépend non seulement de la taille de l'échantillon mais aussi la valeur atteinte par l'indicateur mesuré.

Dans le cas d'une expérimentation randomisée, la situation de départ des membres du groupe bénéficiaire (B) est en moyenne la même que celle du groupe témoin (T) puisque les deux groupes ont été constitués de façon aléatoire. Il s'agit de voir si en fin d'expérimentation leur situation est significativement différente ou non.

L'intervalle de confiance (IC) d'un écart entre deux proportions p_1 et p_2, estimées sur des populations de taille respective n_1 et n_2, est donné par la formule suivante :

$$IC= \widehat{p_1} - \widehat{p_2} \pm 1{,}96\sqrt{\frac{\widehat{p_1}(100-\widehat{p_1})}{n1} + \frac{\widehat{p_2}(100-\widehat{p_2})}{n2}}$$

Cette formule fonctionne dans le cas d'une sélection venant d'un sondage aléatoire simple, ce qui n'est pas toujours le cas. Certaines modalités d'enquête peuvent améliorer la situation présente (tirage stratifié, sondage équilibré, calage) ou la détériorer (tirage à plusieurs degrés, par exemple si l'on tire des zones et des individus dans la zone, on a alors de l'effet de grappe). Cette formule permet néanmoins de fournir un scénario central et donne des ordres de grandeur relativement robustes.

La formule correspond à un intervalle de confiance avec une probabilité de 95%, c'est-à-dire que dans 5 % des cas la différence des proportions réelles sortira de cet intervalle de confiance. Autrement dit, si l'on décide que les deux populations sont différentes à partir du moment où l'IC ne comprend pas la valeur 0, on aura tort dans moins de 5% des cas (cas où 0 est une des bornes de l'IC).

Le tableau 2 donne la valeur de n ($=n_1=n_2$) telle que IC > 0 (dans le cas où $\widehat{p_1} > \widehat{p_2}$).

Tableau 2 : Taille d'échantillon théoriquement requise pour chaque sous-population B et T, selon le degré de précision souhaité et l'ampleur du phénomène analysé

Valeur de l'indicateur mesuré	Degré de précision souhaité					
	1 point	2 points	3 points	4 points	5 points	10 points
5% (ou 95%)	4 000	1 100	550	350	250	100
10% (ou 90%)	> 8 000	2 000	900	550	350	100
15% (ou 85%)	> 8 000	2 700	1 200	700	450	150
20% (ou 80%)	> 8 000	3 200	1 500	850	***550***	150
25% (ou 75%)	> 8 000	3 700	1 700	950	650	200
30% (ou 70%)	> 8 000	> 8 000	1 850	1 050	700	200
35% (ou 65%)	> 8 000	> 8 000	2 000	1 150	750	200
40% (ou 60%)	> 8 000	> 8 000	2 150	1 200	750	200
45% (ou 55%)	> 8 000	> 8 000	2 150	1 200	800	200
50%	> 8 000	> 8 000	2 150	1 200	800	200

Lecture : on peut considérer qu'un écart de 5 points (où l'indicateur analysé vaut 20% (ou 80%) dans l'un des groupes et 25% (ou 85%) dans l'autre) est statistiquement significatif (avec une probabilité de 95%) à partir d'un échantillon total minimum de 550 bénéficiaires et 550 témoins répondants. En dessous de cette taille d'échantillon, cet écart n'est pas significatif. Pour qu'un écart de 2 points soit significatif quand l'indicateur analysé prend des valeurs plus élevées, la taille de l'échantillon doit être augmentée.

Remarques méthodologiques

La précision d'une estimation varie selon que le phénomène analysé est plus ou moins répandu. Le risque de se tromper diminue fortement dès que les bornes de l'IC s'éloignent de 0. Considérer qu'un phénomène est rare revient à considérer que son absence est très répandue, et inversement, considérer qu'un phénomène est très répandu revient à considérer que son absence est rare. Autrement dit, c'est bien la même situation qui est analysée que l'on analyse la proportion p ou la proportion 100-p. Il en découle que la précision d'une estimation est symétrique par rapport à 50%.

A effectif donné, la précision diminue à mesure que la valeur de l'indicateur analysé se rapproche de 50%. C'est pourquoi, pour atteindre un même degré de précision il est nécessaire de se fonder sur des échantillons plus importants pour étudier des phénomènes qui concernent *environ* la moitié d'une population que pour des phénomènes peu répandus.

De petites tailles d'échantillon (100 bénéficiaires et 100 témoins) sont en théorie suffisantes pour mesurer des écarts entre deux populations sur des phénomènes très répandus (ou très rares). Rappelons toutefois qu'ici l'**hypothèse sous-jacente** est que les deux groupes B et T sont constitués aléatoirement et initialement semblables. Or, ce caractère aléatoire n'est assuré qu'à partir d'une certaine taille d'échantillon. En particulier, quand les populations concernées sont très hétérogènes (en termes de niveau de diplôme par exemple), il est souhaitable que les échantillons soient « suffisamment importants ». **Un compromis entre le coût des enquêtes et le respect du caractère aléatoire doit donc être trouvé lors de la détermination d'une taille d'échantillon optimale.**

De plus, si le tirage au sort ne porte pas sur les individus mais par exemple sur des missions locales ou des départements, il importe de sélectionner un grand nombre d'unités pour assurer la comparabilité des populations bénéficiaire et témoin.

Les tailles d'échantillons et les mesures de précisions indiquées dans le tableau ci-dessus ne tiennent pas compte de l'attrition. Elles correspondent à un taux de réponse de 100%. L'attrition diminue la précision des indicateurs (augmente l'intervalle de confiance). Ainsi si l'on prévoit qu'un tiers des individus des deux populations ne répondront pas, il faut dès le départ envisager un échantillon plus grand.

Quelques points de repères

(1) Part des 18-24 ans au chômage : 22%
(2) Part des 18-24 ans en emploi : 30%
(3) Part des 18-29 ans en logement autonome : 56%
(4) Part des 18-24 ans qui ont le permis de conduire : 65%
Sources : (1) et (2) Insee, enquête Emploi 2011 ; (3) Insee, Enquête Nationale Logement, 2006 ; (4) Insee, Enquête Nationale transport 2005

I.1.2. Des solutions alternatives et leurs limites

Une comparaison avec un autre territoire

Afin d'éviter un phénomène de contagion entre bénéficiaires et témoins, l'échantillon témoin peut être choisi sur un autre territoire, si possible comparable au territoire bénéficiaire.

Cette solution avait été proposée dès le départ dans un projet qui cherchait à soutenir les lycéens dans leurs choix d'orientation. Fondé sur le volontariat des professeurs principaux, concernait des classes dans leur ensemble. Il a donc été envisagé dans un premier temps de constituer l'échantillon témoin à partir de classes de même série (S, ES ou L) que les classes expérimentales, dans les mêmes lycées. Cependant, il s'est avéré qu'il existait un risque important de contagion : échanges entre les professeurs, les élèves… De plus, certaines actions pouvaient concerner l'académie entière comme des journées portes-ouvertes ou la mise en place d'un site Internet. Le tirage aléatoire a donc été abandonné et remplacé par deux échantillons témoin :

- Un premier comprenant des classes de même série que celle des lycées bénéficiaires dans des lycées de l'académie ne bénéficiant pas directement du dispositif
- Un second comprenant des classes de même série que celle des lycées bénéficiaires dans des lycées hors académie afin de tenir compte du fait que l'ensemble des lycées de l'académie bénéficiaient, même indirectement, du dispositif.

C'est également le parti pris qui avait été choisi dans le cas des plateformes de lutte contre le décrochage mises en place par deux missions locales. Le projet prévoyait une forte mobilisation des conseillers de la mission locale sur ce public de jeunes décrocheurs, ainsi que l'activation d'un partenariat local. Or, si les jeunes bénéficiaires et témoins étaient choisis sur le même territoire, un risque de contagion des échantillons par le biais des conseillers était possible. Ces derniers auraient pu faire profiter volontairement ou non aux témoins des actions du dispositif. Afin d'éviter cet effet de contagion, deux autres missions locales avaient été choisies pour y constituer un échantillon témoin. Cette solution n'assure pas que les échantillons soient comparables car s'il est possible de contrôler que les deux échantillons ont les mêmes caractéristiques observables, cette méthode ne garantit pas qu'ils aient les mêmes caractéristiques inobservables. Par ailleurs, la lenteur d'intégration des jeunes dans les plateformes semblait indiquer que l'échantillon témoin ne peut être considéré comme comparable. 143 jeunes témoins ont répondu au questionnaire tandis que dans le même temps seuls 48 jeunes avaient intégré les plateformes dans les missions locales expérimentatrices. Il a été décidé en accord avec la Mission d'Animation du Fonds d'expérimentation sur la Jeunesse de ne pas poursuivre cette interrogation de l'échantillon témoin. En effet, il était probable qu'une forte majorité des témoins auraient décliné la proposition d'intégrer une plateforme si une telle proposition leur était faite. C'est pourquoi l'interrogation du groupe témoin et la deuxième interrogation ont été abandonnées.

Dans un projet sur le logement, cette solution de comparaison avec un autre territoire a été choisie afin de pallier le refus, de la part des missions locales, de mettre en place un tirage aléatoire. Dans ce cas, des différences de profils entre les échantillons bénéficiaire et témoin ont été constatées. En effet, pour pouvoir être considérés comme bénéficiaire, les jeunes des Missions locales porteuses du projet devaient remplir un certain nombre de critères et notamment avoir un revenu suffisant pour être capables de financer un loyer. Or, la mission locale du territoire témoin n'a pas appliqué ce critère sur ses jeunes témoins. Il en résulte que ces derniers avaient des revenus plus faibles que les bénéficiaires.

De façon générale il s'avère souvent difficile d'identifier un territoire jumeau pour en faire un territoire témoin. En particulier pour des projets très ancrés localement et marqués par le contexte.

La définition de groupes témoins par comparaison géographique pour palier l'impossibilité de réaliser un tirage aléatoire directement au niveau du public éligible ne doit se faire qu'en sélectionnant **un grand nombre de territoires bénéficiaires et témoins** (au moins une dizaine). Ce n'est qu'ainsi que la comparaison entre « territoires bénéficiaires » et « territoires témoins » est scientifiquement pertinente.

Une comparaison dans le temps

Une autre façon de constituer un groupe témoin est d'interroger un échantillon sur le même territoire que l'échantillon bénéficiaire, ayant connu la même problématique que l'échantillon bénéficiaire mais avant la mise en place de l'expérimentation. Cette méthode a été adoptée dans un projet visant à développer l'accès à l'apprentissage et aux contrats de professionnalisation de jeunes des missions locales. Ce dispositif aspirait à réduire les risques de rupture en mobilisant fortement les conseillers des missions locales pour un suivi renforcé des jeunes et en développant le partenariat avec les Centres de Formation des Apprentis (CFA), les chambres consulaires et les organismes paritaires collecteurs agréés (OPCA).

La définition d'un groupe de témoin dans la même zone géographique mais sur une période antérieure à l'expérimentation a pour principal inconvénient le fait que l'on ne se place pas dans le même contexte socioéconomique. S'il est difficile de s'affranchir des évolutions de contexte, il convient que l'évaluateur soit **attentif aux différences entre les périodes analysées** pour ne pas attribuer au programme expérimenté des effets qui proviennent de changements économiques globaux, et inversement, pour ne pas sous-estimer les effets d'un programme qui se serait déroulé dans un contexte moins favorable. Dans le cas mentionné ci-dessus, l'expérimentation s'est en fait déroulée en même temps que des campagnes nationales d'information en faveur de l'alternance et le financement de postes de développeurs de l'alternance dans le cadre du plan en faveur de l'emploi des jeunes présenté par l'Etat en avril 2009. Il est probable que le projet comme ses bénéficiaires aient été portés par ce contexte.

De plus, dans le cas de comparaisons temporelles entre bénéficiaires et témoins, il faut **s'assurer que la période d'observation des témoins et la période d'exposition au programme des bénéficiaires soient de même durée** afin de comparer les parcours avec le même recul.

Dans le cas précis de ce projet sur l'alternance, les délais de mise en place de la phase d'enquête (en particulier ceux de récupération des listes de jeunes) ont décalé la première vague d'enquête. Par conséquent la période de suivi des bénéficiaires a été plus courte que la période de référence sur lesquels les témoins étaient interrogés. Cet écart semble avoir eu peu d'incidence sur les réponses concernant les parcours antérieurs à l'alternance, les motivations à intégrer l'alternance et les soutiens reçus dans cette étape. En revanche, cette différence de durée rend impossible la comparaison entre les situations après le déroulement du contrat entre bénéficiaires et témoins. Une autre solution aurait été d'interroger uniquement parmi les bénéficiaires ceux qui avaient intégré une formation en alternance en 2010, mais les effectifs à interroger auraient été trop faibles.

Les limites des alternatives à mieux prendre en compte dès la conception

Des méthodes alternatives à la randomisation ont ainsi été déployées dans de nombreux projets. Par rapport à la constitution aléatoire d'échantillons comparables, toutes présentent des inconvénients qui les rendent moins performantes. Dans le cas des expérimentations pour la jeunesse, les méthodes alternatives ont rarement permis de mesurer l'effet propre des dispositifs car malgré tout il subsistait trop de dissemblances entre les échantillons témoins et bénéficiaires. Ces constats ont souvent conduit à abandonner la comparaison entre les deux publics en raison des biais de sélection doublés de faibles effectifs (qui ne permettent pas la mise en œuvre de méthodes économétriques pour neutraliser les biais).

Remarquons que ces abandons se sont souvent faits en plusieurs étapes. Dans un premier temps, l'échantillon randomisé a été remplacé par un groupe témoin constitué de façon raisonnée (jeunes ayant des caractéristiques semblables, sur le même territoire ou sur un territoire voisin). Mais la faiblesse des effectifs, les faibles taux de réponse de l'échantillon témoin dans le temps, les différences de profil ou entre les périodes d'interrogation entre bénéficiaires et témoins ont souvent amené à renoncer à poursuivre la collecte ou l'analyse des réponses des échantillons témoins.

Finalement, parmi les projets évalués par le CRÉDOC :

- 3 évaluations comprenaient la mise en place et le suivi d'un échantillon témoin aléatoire bénéficiaire et témoin.
- 7 évaluations comprenaient un échantillon témoin raisonné, même si pour 3 d'entre elles, les vagues d'enquêtes auprès des témoins n'ont pas toutes été effectuées (témoins trop différents des bénéficiaires ou impossibles à recontacter).
- 3 évaluations ont du abandonné l'échantillon témoin prévu (aléatoire ou raisonné).
- 4 évaluations ne comprenaient pas d'échantillon témoin dans leur protocole initial.

En fin de compte, afin de pallier la perte d'informations induite par la suppression des échantillons témoins, des entretiens qualitatifs (cf. partie sur les méthodes qualitatives), collectifs et individuels, ont été mis en place auprès des jeunes du public cible. Ces entretiens ont permis alors de mieux connaitre le parcours des jeunes, leurs attentes et leurs opinions sur le dispositif.

I.2. Les jeunes : un public peu accessible dans les enquêtes quantitatives

La quasi-totalité des évaluations du CRÉDOC pour le Fonds d'expérimentation pour la Jeunesse comprenaient deux à trois interrogations quantitatives auprès des bénéficiaires et des témoins : au début de l'expérimentation, à la fin et parfois à mi-parcours. Ces interrogations visaient à connaitre la situation et le parcours du jeune. La plupart des interrogations ont été réalisées par téléphone. Certains questionnaires ont néanmoins été administrés en face à face, notamment pour les premières interrogations de jeunes concernés par les projets main de justice.

I.2.1. Une forte attrition entre la première et la seconde interrogation

Lors de la première vague d'enquête, l'objectif en termes de nombre de jeunes interrogés est atteint pour la moitié des expérimentations. Les premières enquêtes des autres projets atteignent entre les deux tiers et 85% des objectifs fixés. Le taux d'attrition entre la première et la seconde enquête devient problématique. Indépendamment du taux d'attrition, les échantillons de départ étaient faibles sur certains projets, d'où un nombre de répondants faible en deuxième interrogation. Par exemple, les quatre projets sur la thématique « main de justice » portaient sur des publics comptant en général autour de 70 à 80 bénéficiaires. Si à la première interrogation, entre 50 et 70 jeunes ont pu être interrogés, ce qui constitue un taux de réponse correct, à la seconde interrogation environ 20 jeunes seulement par projet ont répondu aux questionnaires. Une partie de ces jeunes étaient encore en détention (soit jamais libérés, soit retournés en détention), d'autres étaient impossibles à joindre, une petite minorité d'entre eux refusant explicitement d'être réinterrogés. Sur ces quatre projets « main de justice », les troisièmes interrogations prévues en fin d'expérimentation ont par conséquent été supprimées et remplacées par des entretiens qualitatifs auprès des jeunes bénéficiaires.

De façon générale, l'attrition entraine plusieurs biais :

- **un risque classique de biais de représentativité par rapport à l'échantillon initial** : les jeunes finalement interrogés en dernière vague représentent une proportion trop faible de la population mère pour être représentatifs. La validité interne de l'évaluation est alors remise en cause.
- **Un risque de biais lié à une attrition « différentielle »** : une attrition « différentielle » entre les deux échantillons bénéficiaire et témoin c'est-à-dire plus importante pour un groupe que pour l'autre, entrainerait un biais dans la comparaison des échantillons. Même si en apparence, notamment dans le cas de la mise en place d'un redressement sur les échantillons, les deux groupes semblent identiques en comparant leurs caractéristiques observables, il est impossible d'en conclure qu'ils le sont quant à leurs caractéristiques inobservables. Les différences observées ne peuvent alors systématiquement être imputées au dispositif.

Ce phénomène d'attrition a été rencontré sur toutes les thématiques : insertion des jeunes sous main de justice, accès à un logement, accès à la formation par alternance et décrochage scolaire. La seule exception concerne un projet sur l'orientation des lycéens. Il s'agit d'un public plus jeune en moyenne, moins mobile, moins fragile, plus proche de ses parents, donc plus facilement joignable.

L'attrition a souvent été plus importante pour les témoins que pour les bénéficiaires. Les bénéficiaires, du fait de leur suivi par les porteurs, ont eu plus de contacts avec leurs conseillers. Leurs coordonnées ont donc été plus souvent mises à jour que celles des témoins.

I.2.2. Les spécificités du public « jeunes »

Notre expérience nous confirme que les jeunes dans leur ensemble sont un public plus difficile à joindre et à suivre, et ce pour plusieurs raisons :

- **Un public très mobile** : les jeunes sont plus souvent en recherche d'emploi, de logement ou de formation et sont donc par définition un public très mobile, changeant souvent d'adresse et de numéro de téléphone fixe. Si les jeunes ne prennent pas contact avec les porteurs et si ces derniers ne peuvent pas mettre à jour leurs coordonnées, il devient extrêmement difficile de les contacter.
- **Les numéros de portables invalides** : les numéros de téléphone disponibles correspondaient souvent à un numéro de portable. Or, les jeunes ont souvent recours aux cartes prépayées, changent régulièrement de téléphone et de numéro (ce dernier phénomène ayant tendance à se réduire depuis qu'il est possible de changer d'opérateur tout en conservant son numéro). Quand eux-mêmes n'ont pas de téléphone, ils ont souvent indiqué ceux de leurs parents ou d'amis qui au moment de l'enquête ne connaissent pas toujours le moyen de contacter le jeune à interroger.
- **Un public qui ne se sent pas toujours concerné** : les bénéficiaires n'ont pas systématiquement été mis au courant qu'ils participaient à une évaluation, ce qui peut expliquer une partie des refus de répondre. Parfois, bien qu'ils aient été identifiés comme ayant bénéficié d'une aide particulière, ils estiment celle-ci peu ou pas utile ou faisant partie d'un ensemble d'aides possibles. Ils se montrent alors désintéressés et refusent de répondre au questionnaire. Par exemple, sur une deuxième interrogation d'un projet sur la thématique logement, bien que peu nombreux, les refus pour cette raison représentent 3% des non-répondants. Les témoins quant à eux ne perçoivent souvent pas l'intérêt de répondre aux enquêtes.
- **Un public très hétérogène** : entre 18 et 25 ans les jeunes connaissent des situations très variées. Même à âge donné les niveaux de diplôme sont très épars. Entre étude et emploi, les statuts sont nombreux et changent vite, en particulier pour des publics qui connaissent des situations précaires. Sur la durée d'une expérimentation, un jeune a pu avoir plusieurs CDD ou missions d'intérim, alterner des périodes de chômage et d'emploi, des formations plus ou moins

courtes ou interrompues. Sur le plan personnel également, les situations familiales évoluent rapidement, surtout quand le lien avec les parents est distendu. Comme cela a été mentionné plus haut, c'est au cours de ces premières années de vie adulte que les déménagements sont fréquents et la question du logement centrale.

I.2.3. Les enquêtes par questionnaire : cibler les indicateurs clés

Les enquêtes par questionnaire visaient à collecter des informations homogènes sur les jeunes bénéficiaires et témoins (situation professionnelle, leur situation de logement…), à différentes dates afin de comparer l'évolution de leur situation.

Les questionnaires étaient prévus pour durer en moyenne une vingtaine de minutes afin de collecter des informations les plus précises possibles sur les jeunes et appréhender la variété des situations rencontrées et approfondir des cas particuliers. Sur ce plan, il est probable que les questionnaires conçus ont été généralement trop ambitieux. Ils ont apparemment été perçus comme trop longs par les jeunes interrogés. Les questionnaires un peu longs en première vague d'enquête ont vraisemblablement découragé les jeunes à répondre plus tard lors de la deuxième interrogation.

En termes d'exploitation, les questions spécifiques aux différentes situations possibles ont concerné finalement des effectifs faibles, conduisant par conséquent à une information inexploitable. Un exemple dans le cadre d'un dispositif cherchant à favoriser l'accès au logement : pour cette expérimentation, un échantillon bénéficiaire et un échantillon témoin ont été créés de façon aléatoire. Les échantillons ainsi obtenus étaient composés de 700 jeunes environ chacun, soit un effectif « confortable » de 1 400 jeunes au total. Un peu moins de 900 jeunes ont été interrogés lors de la première vague, dont environ la moitié étaient bénéficiaires et la moitié témoins. Parmi eux, 300 environ étaient locataires. De nombreuses questions portent uniquement sur ces jeunes, notamment des questions sur les impayés de loyer et de charges. Or, seuls 77 locataires, bénéficiaires comme témoins, ont déclaré avoir eu des impayés dans le logement qu'ils occupaient au moment de l'interrogation. Il est alors difficile de faire des statistiques sur le profil de ces jeunes d'effectif restreint. De même pour les questions portant sur les raisons du départ d'un précédent logement en location : environ, 180 jeunes seulement parmi les 900 avaient été locataires précédemment et 12 seulement sont concernés par les questions portant sur les raisons pour lesquels le propriétaire leur a demandé de quitter le logement. D'autres questions sont donc inexploitables.

Enfin, le risque à utiliser des questionnaires trop complexes, comportant des questions abstraites ou nécessitant d'avoir un recul sur sa situation, est de perdre des répondants en cours d'interrogation ou bien d'être confronté au refus de réponses à certaines de ces questions.

Nos recommandations sur les questionnaires sont par conséquent les suivantes :

- des questions simples et ciblées, compréhensibles par tous, de façon à obtenir sur des indicateurs précis sur la situation au moment de l'enquête et l'opinion quant au dispositif intégré ;
- pas trop de filtres qui risquent de réduire les échantillons et rendent certaines questions inexploitables ;
- une durée de 10 à 15 minutes maximum pour éviter les risques d'abandon en cours d'interrogation et dans les interrogations suivantes.

I.3. Les tableaux de bord : des outils précieux qui nécessitent un remplissage harmonisé

L'objectif de ces tableaux de bord est de collecter, pour chaque bénéficiaire, de façon harmonisée, sa situation à l'entrée dans le dispositif, le nombre et la nature des actions suivies, la situation à chaque contact avec un conseiller et la situation à la fin du suivi. Mis en place dans la totalité des protocoles d'évaluation ils ont permis de compléter les interrogations quantitatives.

Deux types de tableaux de bord ont été utilisés :

- des extractions de logiciel de gestion (de type Parcours 3 pour les Missions locales par exemple) ;
- des tableaux de bord *ad hoc* mis en place à la demande de l'évaluateur mais en concertation avec le porteur de projet qui devait le remplir. L'avantage de ces tableaux est de pouvoir intégrer des informations qui n'existaient pas dans les logiciels de gestion voire palier l'absence de logiciel de gestion. Dans le cas des deux projets nationaux notamment, les tableaux de bord ont constitué un outil commun pour des acteurs n'appartenant pas tous aux mêmes structures :
    - pour Agiloje : des missions locales, des Foyers Jeunes Travailleurs (FJT), des associations, des Bureaux Informations Jeunesse (BIJ), des Centres Communaux d'Action Sociale (CCAS), des organismes PACT pour le logement…
    - pour le projet d'apprentissage porté par le CNML : des missions locales, des CFA.
    - pour la « Plateforme logements Jeunes 94 » : les missions locales du Val-de-Marne et le Conseil Général.

Dans le cas des deux premiers, le tableau de bord a pris la forme d'un Extranet permettant :

- la saisie du nom et des coordonnées des jeunes pour les enquêtes quantitatives ;
- le tirage aléatoire par l'extranet des jeunes et l'affectation dans les groupes bénéficiaire et témoin sans l'intervention des acteurs ;
- le suivi de la situation des jeunes.

Ces tableaux de bord doivent être remplis régulièrement et de façon la plus homogène possible. Il est important de noter que le remplissage de ces tableaux constitue en fait une contrainte supplémentaire pour les porteurs de projet et les acteurs, notamment pour les tableaux de bord *ad hoc* qui ne font pas partie des outils de suivi usuels. En outre, dans certains cas cette saisie constitue une double saisie avec le logiciel de gestion existant. Ne pas alourdir la charge des acteurs est un aspect important à prendre en compte par les porteurs comme par les évaluateurs lors de la conception de ces outils.

Comme cela a été mentionné plus haut, ces tableaux ont utilement complété les informations tirées des enquêtes et apportaient des informations qu'il n'était pas possible de collecter autrement sans alourdir encore le questionnement ou être confronté à des problèmes de mémoire : fréquence et contenu des contacts avec les acteurs de l'accompagnement, situation de logement ou vis-à-vis de l'emploi à différentes périodes. L'exploitation des tableaux de bord a ainsi permis le calcul d'indicateurs tels que le nombre de rendez-vous avec les missions locales, le nombre de périodes de chômage, la durée des formations. Ces indicateurs ont permis de mesurer la mise en œuvre des dispositifs et de mieux connaitre les populations suivies. En particulier, les tableaux de bord ont fourni une information exhaustive sur la population bénéficiaire car ils étaient généralement remplis pour l'ensemble des bénéficiaires tandis que les enquêtes portaient sur un échantillon de bénéficiaires (notamment dans les quelques projets de grande ampleur).

Pour ces expérimentations de l'AP1, sauf dans quelques cas, les tableaux de bord *ad hoc* ne suivaient d'ailleurs que les jeunes bénéficiaires. Il aurait été intéressant de renseigner de la même façon les actions de droit commun suivies par les témoins et leur situation à chaque contact avec le porteur. Ceci aurait alors permis de comparer les situations des témoins à celles des bénéficiaires en se fondant sur des informations provenant de la même source et d'obtenir des informations sur les témoins qui ont souvent manqué dans le recueil des données par questionnaire.

Les autres points d'attention lors de la conception de tableaux de bord de suivi renvoient aux difficultés que nous avons relevées lors de la complétion par les acteurs de terrain et lors de l'exploitation :

- **Eviter des tableaux complexes**. Ces derniers contiennent souvent plusieurs dimensions - une entrée par bénéficiaire, une entrée par date et parfois une entrée par actions/actualités – et ce sur plusieurs fichiers ou onglets différents. L'évaluateur doit participer à l'élaboration du tableau de bord, en listant les informations qu'il juge nécessaire, et, plus encore que pour les enquêtes quantitatives, anticiper au maximum la phase de traitement de ces tableaux. Dans de nombreux cas, aucun format particulier de fichier n'avait été spécifié. Or les acteurs se sont appropriés les tableaux Excel de façons différentes ce qui a rendu l'exploitation complexe. Nous recommandons aux évaluateurs de spécifier strictement la forme des tableaux, par exemple en consacrant une seule ligne par jeune, ou un seul onglet pré-formaté, et définir le contenu de chaque fichier, et ce toujours en lien avec les futurs utilisateurs.

- **Remplir les tableaux de façon homogène**. Quand plusieurs acteurs sont en charge de la mise à jour du tableau de bord, tous n'ont pas la même façon de le remplir. Par exemple, pour un projet multi-sites, certains acteurs ont noté tous les contacts avec les jeunes (rendez-vous en face à face, appels téléphoniques, SMS ou mails) tandis que d'autres ne vont avoir inscrit dans le tableau de bord que les rendez-vous en face à face ou les appels téléphoniques les plus importants. Le nombre moyen d'actions variait alors fortement d'un territoire à l'autre. De plus, les sites n'ayant pas tous mis en place les mêmes actions, les items portant sur ces actions n'ont pas été suffisamment bien définis, ni suffisamment bien arrêtés. Au final, le tableau de bord comprenait un nombre d'actions trop important, pour lesquelles il était compliqué de savoir à quoi elles correspondaient. Cette information n'a donc pu être pleinement utilisée. Enfin, il existe toujours un doute sur la qualité de remplissage du tableau de bord : les porteurs ont-ils toujours inscrit toutes les informations disponibles ou n'ont-ils pu le faire par manque de temps, de moyens.

- **Utiliser les tableaux tout le temps de l'évaluation.** Quand l'expérimentation se termine, les acteurs ne saisissent plus les informations dans les tableaux de bord. Le temps de l'évaluation n'étant pas toujours celui de l'expérimentation, l'évaluateur n'a parfois plus disposé d'informations récentes sur les bénéficiaires. Il convient donc de faire coïncider la période de remplissage des tableaux avec celle de l'évaluation.

II. DES APPROCHES QUALITATIVES A DEVELOPPER DANS LES EXPÉRIMENTATIONS

Les entretiens avec les acteurs (porteurs de projet, partenaires) ont pour principal objectif de rendre compte des conditions pratiques de mise en œuvre des expérimentations. Ils fournissent des informations sur les spécificités du territoire de l'expérimentation. Ils sont l'occasion de prendre en compte les retours d'expérience des différents acteurs concernés par les expérimentations. Ils donnent des indications précieuses pour réfléchir aux conditions de généralisation ou d'essaimage sur d'autres territoires. Enfin, il est important de croiser les points de vue des différents partenaires avec ceux des porteurs, des jeunes et avec les résultats quantitatifs afin d'obtenir l'information la plus complète et la plus objective possible.

II.1. Les entretiens auprès des porteurs pour éclairer les conditions de mise en œuvre des actions

Des entretiens ont eu lieu avec les porteurs tout au long de l'évaluation. Les entretiens en début d'expérimentation ont permis de connaitre les pratiques antérieures du porteur par rapport au public concerné par l'évaluation, les conditions de développement du projet, les objectifs visés et l'état du partenariat. Dans les cas où les projets ont été mis en œuvre directement par la structure porteuse du projet, les entretiens ont été menés à la fois avec les responsables des projets et avec les personnels directement opérationnels qui étaient en contact direct avec les jeunes.

L'évaluateur a maintenu les contacts tout au long des expérimentations, permettant d'analyser les évolutions par rapport au projet initial, les difficultés éventuelles rencontrées et les adaptations alors nécessaires dans les méthodes.

En fin d'expérimentation, les entretiens avec les porteurs ont permis de recueillir le bilan fait par les acteurs sur les actions réellement mises en œuvre, leur point de vue concernant les effets sur les bénéficiaires ainsi que sur les évolutions des relations partenariales.

Il est important de croiser le point de vue des porteurs avec celui des partenaires ainsi qu'avec les résultats des enquêtes quantitatives ou qualitatives sur les bénéficiaires notamment quand les porteurs de projet, parfois dans l'optique de pérenniser les actions quels que soient les résultats de l'évaluation, passent sous silence certaines des difficultés rencontrées.

II.2. Les entretiens auprès des partenaires pour mettre en évidence les synergies et les freins

Les dispositifs expérimentés reposent le plus souvent sur une synergie des acteurs. La consultation des partenaires donne l'opportunité de faire un diagnostic de cette synergie attendue.

Sortir du point de vue du seul porteur de projet

Des entretiens systématiques avec les autres partenaires du projet en début et en fin d'expérimentation permettent d'élargir le point de vue du porteur de projet, de vérifier quels sont les points d'accord et les différences d'analyse sur les conséquences de l'expérimentation. A l'étape de lancement du programme, ils sont l'occasion de compléter le panorama de l'état des partenariats existants et des actions déjà en place sur le territoire. Ce bilan est particulièrement important à réaliser pour mesurer en fin de programme le chemin parcouru. Ces entretiens mettent parfois en évidence des différences dans les attentes vis-à-vis du programme qui peuvent expliquer des manques de synergie dans le partenariat. Ils complètent la connaissance de l'évaluateur sur les interventions des différents acteurs du programme.

Ces entretiens ont pu parfois mettre en lumière l'absence de visibilité d'un programme pour les participants, alors que le porteur pouvait penser que certains partenaires avaient bien identifié les actions mises en place.

Le risque d'une liste incomplète de partenaires

Dans un premier temps, l'évaluateur part généralement d'une liste de partenaires qui lui a été communiquée par le porteur. Il s'agit souvent du « premier cercle » des partenaires, de ceux dont le porteur est le plus proche. Le risque existe d'une sélection des partenaires partageant le point de vue du porteur sur le bilan de l'expérimentation. Le CRÉDOC a généralement demandé aux acteurs interviewés de suggérer une liste d'autres partenaires à interviewer. Ainsi le champ des personnes prises en compte dans les phases qualitatives peut être élargi.

Dans le cas de projets visant à l'insertion professionnelle, le plus difficile a souvent été d'obtenir les contacts des employeurs (chefs de petites entreprises ou responsables de personnel) avec qui un partenariat s'était noué, particulièrement quand ces employeurs ne relèvent pas de l'insertion par l'activité économique. Parfois le porteur de projet craignait que les employeurs –qui sont des personnes difficiles à mobiliser – ne se sentent importunés par la demande d'entretiens par l'évaluateur et se montrent moins disponibles en cas de sollicitations d'autres intégrations dans leur entreprise de personnes suivies par leur structure. Dans le cas d'un des projets concernant des jeunes sous main de justice, les entreprises participent à un programme d'insertion d'anciens détenus sans connaitre le passé des candidats qui leur étaient envoyés par l'organisme d'insertion. Dans ce cas il n'y aurait donc pas eu de sens de les interroger sur ce programme dont ils ignoraient l'existence.

Bien identifier les catégories de partenaires à interroger

Pour s'assurer que l'analyse qualitative des partenaires est complète, l'évaluateur doit identifier les catégories de partenaires pour chaque expérimentation. Ces catégories sont spécifiques aux thématiques du projet. Pratiquement toujours, les représentants de collectivités territoriales sont à rencontrer (communes, agglomérations, conseils généraux ou régionaux…). Tous les projets ayant une dimension d'insertion professionnelle supposent de comprendre comment se fait l'articulation avec les services de Pôle Emploi. Dans le cas de projets concernant les jeunes sous main de justice, les juges d'application des peines et les services d'insertion et de probation sont des partenaires incontournables à prendre en compte. Pour l'alternance, centres de formation, chambres consulaires et organismes paritaires collecteurs agréés sont à interviewer. Les partenaires dans le champ du logement vont des gestionnaires de foyers aux services logement de collectivités.

II.3. Les entretiens auprès des jeunes pour recueillir des récits de parcours et leur perception des dispositifs

Les projets d'évaluation du CRÉDOC prévoyaient la réalisation d'enquêtes quantitatives répétées auprès des jeunes. Le fort taux d'attrition n'a pas toujours permis de réinterroger un échantillon suffisant de jeunes. Dans ces cas, la substitution s'est souvent faite par une campagne d'entretiens qualitatifs auprès des jeunes qu'il était encore possible de joindre.

Les réponses aux premières vagues de questionnaires permettent de connaitre certains des traits particuliers des jeunes qui ont été interviewés par entretiens approfondis, par rapport à l'ensemble des bénéficiaires de l'expérimentation. Il reste possible qu'ils présentent des caractéristiques particulières non identifiables à partir des informations recueillies au départ (plus grande motivation dans l'expérimentation par exemple, ou plus faible mobilité géographique depuis le premier questionnaire). C'est pourquoi une certaine prudence s'impose sur leur représentativité par rapport à l'ensemble du public touché par l'expérimentation.

Le ressenti du programme par les jeunes

A travers le récit que les jeunes font de leur parcours depuis leur entrée dans le programme, il est possible d'avoir accès au point de vue des bénéficiaires des actions sur les effets propres de l'expérimentation. Leur ressenti peut être approfondi grâce à ces étapes qualitatives. Les avancées en termes de parcours peuvent être ou non considérées par les jeunes comme une conséquence de leur participation au dispositif. Ces entretiens apportent un point de vue complémentaire à celui des acteurs institutionnels sur les effets propres des actions.

Les entretiens de bénéficiaires menées par un évaluateur extérieur donnent l'opportunité d'aborder la question de l'appropriation du dispositif par son public et sa compréhension. Les jeunes peuvent

apporter leur point de vue sur l'intérêt du dispositif et faire des suggestions sur des améliorations possibles en cas de pérennisation ou d'essaimage.

PARTIE 3 : REFLEXIONS SUR LES CONDITIONS DE REUSSITE DES EXPERIMENTATIONS ET DE LEUR EVALUATION

Du côté des évaluateurs, l'abandon des échantillons témoins dans la majorité des dispositifs d'évaluation et la forte attrition ont limité la puissance d'une partie des méthodologies déployées. Du côté des porteurs de projets, la plupart souhaitaient pouvoir communiquer rapidement les résultats de l'évaluation et s'y référer pour pérenniser les actions expérimentées. Ils avaient notamment de fortes attentes en termes de résultats chiffrés. Or, nombre d'expérimentations ne se prêtaient pas à l'emploi de méthodes quantitatives et se sont en outre heurtées à diverses contraintes de terrain. La fin des expérimentations et la rédaction des rapports finaux des évaluations ont ainsi été accompagnées d'un sentiment de frustration de la part des équipes évaluatrices du CRÉDOC comme des porteurs de projets. Par ailleurs, certains des projets expérimentés ont été pérennisés, du moins prolongés, parfois même avant la fin de l'évaluation.

Ces constats invitent à s'interroger d'une part sur la nature des résultats qui peuvent être obtenus d'une expérimentation et, d'autre part, sur la portée et l'utilisation de ces résultats. La question sous-jacente abordée est en fin de compte celle des suites concrètes des expérimentations soutenues par le FEJ, c'est-à-dire celle des leçons à tirer de ces évaluations en termes de politiques pour la jeunesse.

I. LES CONDITIONS DE REUSSITE D'UNE EXPERIMENTATION

Dans sa conclusion, le guide méthodologique de la MAFEJ philosophe : « S'il pense évaluer une expérimentation, en réalité il [l'évaluateur] expérimente une évaluation. » C'est effectivement le constat du CRÉDOC. Nombre de projets ont rencontré des obstacles dans leur déroulement et subi des ajustements en cours d'expérimentations. Malgré ces difficultés très diverses, au final, tous ont pu être évalués. Tirant de notre expérience un certain nombre de clés pour une évaluation réussie, nous revenons ici sur les principales questions qui se sont posées au moment d'analyser les résultats.

I.1. Une montée en charge anticipée pour évaluer les effets dès que les actions sont en place

Certains projets ont connu une montée en charge plus longue que prévue. Une conséquence a été dans certains cas le décalage de la date des premières enquêtes quantitatives pour qu'elles puissent porter sur un échantillon suffisamment grand statistiquement (au moins 100 personnes).

Au moment de la première interrogation, la cohorte de jeunes interrogés était alors composée de jeunes récemment entrés dans le dispositif mais aussi de jeunes présents depuis plusieurs mois.

Inversement, un projet d'accompagnement vers le logement a démarré et porté ses fruits très rapidement : le temps que la première vague d'enquête auprès des jeunes se mette en place (réception et vérification des coordonnées des jeunes, courriers d'information aux jeunes selon les recommandations de la CNIL…), environ un mois et demi après leurs premiers contacts avec le service médiation logement, un grand nombre de jeunes avait bénéficié d'une solution de logement grâce au dispositif. Il en est ressorti une faible évolution de la situation de logement des jeunes entre la première et la deuxième vague d'interrogation, puisque les effets du programme s'étaient fait sentir plus tôt. Les effets du dispositif ont ainsi été sous-estimés dans les enquêtes statistiques Ce sont les éléments tirés des méthodes qualitatives qui ont permis de comprendre la situation.

Ainsi, il s'agit de démarrer l'évaluation au moment approprié, ni trop tôt ni trop tard, en fonction de la temporalité du projet.

I.2. Un public bien ciblé et clairement identifié

L'estimation d'effets propres d'un dispositif sur un public ciblé nécessite que la population bénéficiaire soit définie en lien avec l'objectif de l'expérimentation et bien identifiée.

Les ajustements méthodologiques intervenus en début ou en cours d'expérimentation ont parfois conduit à l'élargissement du public bénéficiaire (en vue d'atteindre des effectifs correspondant à ceux fixés avec la MAFEJ et pour répondre aux exigences de la statistique). Par exemple, dans un projet en faveur des jeunes sous main de justice, le public bénéficiaire initialement composé de jeunes incarcérés et candidats à un aménagement de peine a été élargi à des jeunes en milieu ouvert, c'est-à-dire bénéficiant déjà d'un aménagement de peine (ou condamnés à une peine alternative à l'incarcération). Dans le cas d'un projet visant des décrocheurs universitaires inconnus de la mission locale, la difficulté à mobiliser des listes provenant de l'université d'étudiants en cours d'abandon de leur cursus a amené à intégrer des jeunes déjà suivis par la mission locale impliquée. Un des projets portant sur l'accès au logement a intégré pour partie des jeunes pour lesquels la question de l'accès à un logement autonome n'était pas la priorité.

Ces élargissements ont pour conséquence de brouiller les résultats de l'expérimentation. Autrement dit, quand le public bénéficiaire a été élargi et ne correspond plus au public ciblé dans les objectifs initiaux, on n'observe plus qu'un impact dilué sur une population bénéficiaire plus large. Il n'est alors pas possible de conclure sur l'effet du programme sur le public ciblé initialement.

I.3. Des actions qui diffèrent clairement du droit commun

Plusieurs projets évalués par le CRÉDOC consistaient en fait en un renforcement du droit commun. Du côté des partenariats, cela a impliqué des nouvelles modalités de fonctionnement. Dans le cadre de projets associant missions locales et services pénitentiaires de probation et d'insertion (SPIP), cela s'est par exemple traduit par la mise à disposition, au sein du SPIP, d'une permanence pour un Conseiller Justice de Mission locale qui était ainsi plus souvent présent à la maison d'arrêt, renforçant dans le même temps les échanges et la collaboration entre la Mission locale et le SPIP.

Pour les bénéficiaires en revanche, ces innovations de fonctionnement ne se sont pas forcément matérialisées par une prise en charge plus intensive que pour les jeunes témoins, ou encore que pour les usagers d'un territoire témoin tout aussi dynamique. Aussi, l'évaluateur a-t-il eu des difficultés à mesurer des effets propres sur les bénéficiaires. Le « traitement » (pour reprendre le vocable des économètres) du groupe bénéficiaire doit être différent de celui du groupe témoin sans quoi les différences potentielles de trajectoires entre les deux groupes ne sont pas forcément visibles, et en outre, ne peuvent en aucun cas lui être imputables.

De plus, si les jeunes bénéficiaires sont déjà connus des missions locales, il est difficile d'identifier la plus-value des dispositifs par rapport aux actions antérieures. Autrement dit, dans ces cas, il est difficile de distinguer dans l'évolution de leur situation ce qui est dû à l'accompagnement spécifique de ce qui serait lié au suivi antérieur. Pour déceler les effets propres d'un programme, il est donc préférable que les jeunes qui entrent dans l'expérimentation, en tant que bénéficiaires ou témoins, n'aient pas auparavant été suivis par les structures expérimentatrices - ou alors les jeunes déjà connus de la structure doivent être aléatoirement répartis dans les groupes bénéficiaires ou témoins.

Enfin, l'évaluateur doit avoir à l'esprit ce que recouvre le droit commun. Dans certains territoires, il englobe déjà un grand nombre de mesures en place. Il est alors plus difficile d'identifier les effets des dispositifs expérimentés qui se distinguent peu du droit commun. Par exemple, dans certaines zones d'actions de missions locales où il existait déjà des bourses au logement faisant concorder la demande et l'offre, des partenariats et un accompagnement personnalisé des jeunes, la plus-value du renforcement de ce droit commun est moindre que dans des territoires dépourvus de ces modes d'actions. A moins d'augmenter la taille des échantillons de jeunes faisant partie des expérimentations, les effets marginaux sont alors plus difficiles à déceler (voir encadré 3).

Il reste que, quand l'expérimentation consiste en un renforcement du droit commun identifiable et mesurable, et quand une population témoin a bien pu être définie et suivie, il est alors possible d'observer des effets sur les bénéficiaires comme dans le cas du projet d'orientation de l'université à l'entreprise. Les bénéficiaires ont effectivement eu plus de contacts et un suivi plus rapproché avec l'association porteuse du projet. Bien que les abandons d'études n'aient pu être évités, la situation en emploi est relativement meilleure pour les bénéficiaires que pour les témoins.

I.4. Des actions qui répondent à un besoin identifié et précis

Une autre condition de la réussite d'une évaluation tient dans le fait que les dispositifs expérimentés répondent à un besoin précisément identifié. S'il est pertinent d'aider les jeunes à accéder à un logement autonome, il est apparu que leur proposer un accompagnement dans la durée pour trouver un logement ne répond pas à leurs attentes réelles, car les besoins en logement correspondent à des demandes urgentes. Par exemple, dans le cas de l'accompagnement vers le logement, après un premier contact et des éventuelles visites de logement, le suivi par les professionnels dédiés au public bénéficiaire consistait en des rappels téléphoniques dans les mois suivants, mais restés sans écho du côté des jeunes qui n'en avaient en fait pas besoin. Il est apparu qu'ils se sont rapidement tournés vers leurs proches, le réseau relationnel s'étant montré plus direct et plus efficace. En outre, il semblerait que l'accompagnement proposé dans l'accès au logement autonome peut être contradictoire avec la recherche d'autonomie des jeunes.

De plus, malgré les réflexions menées en amont, certains des projets que le CRÉDOC a évalués n'étaient pas assez précisément définis au moment où l'expérimentation s'est mise en place. Les effets recherchés étaient diffus et les objectifs mal définis malgré une formalisation déjà poussée dans la phase d'élaboration du projet pour candidater au FEJ.

Implicitement, c'est la question de l'adhésion des bénéficiaires au dispositif qui se pose et doit être anticipée. Conformément à la trame de rapport élaborée par la MAFEJ, l'évaluateur était invité à porter « dans la mesure du possible un jugement sur le niveau d'adhésion du public au dispositif et sur de possibles décalages entre le public visé initialement et celui qui a effectivement adhéré au dispositif », et à préciser notamment « si selon lui le dispositif a été suffisamment attractif pour les bénéficiaires ou pas ». L'évaluateur devait, a posteriori donner des précisions sur « les conditions d'éligibilité des bénéficiaires et les modes de prescription » : « ces derniers étaient-ils adaptés ? Les prescripteurs ont-ils adhéré à la philosophie et aux modalités du dispositif ? » S'il est légitime de s'interroger a posteriori sur la façon dont le dispositif a été investi par le public ciblé[37], la question de la pertinence de l'action au regard des critères d'éligibilité doit, d'après notre expérience, être abordée en profondeur dès la conception de l'expérimentation.

I.5. Des financements qui coïncident avec la période d'évaluation

Les protocoles d'évaluation déployés par les évaluateurs comportaient des phases d'entretiens avec les partenaires ainsi que l'interrogation des bénéficiaires (et des témoins), nécessairement programmées en fin d'expérimentation pour en saisir les effets à moyen terme. Il n'était ainsi pas possible que les résultats soient disponibles au moment où les financements s'arrêtent. A fortiori,

[37] Dans certains cas, les bénéficiaires ne sont pas conscients de bénéficier d'un dispositif nouveau ou spécifique.

les résultats ne pouvaient être disponibles avant la fin de l'expérimentation. En outre, trois mois étaient ensuite dévolus à la rédaction du rapport d'évaluation[38]. De plus, les résultats des évaluations étaient en fait destinés à la MAFEJ et devaient d'abord être validés par la mission, qui réservait un mois à cette dernière étape. Sans remarques de la MAFEJ dans ce délai, c'est au bout d'un mois que les rapports des évaluateurs ont été rendus publics et transmis aux porteurs.

Le calendrier des évaluations formaté par la MAFEJ ne coïncidait donc pas complètement avec celui des expérimentations. En outre, les évaluations s'étalaient sur une durée un peu plus longue que celle des financements provenant de la MAFEJ. En pratique, ce décalage a posé de réels problèmes de coordination de l'évaluation et a pu affecter les conditions mêmes de l'évaluation.

Concrètement, pour ces professionnels, fortement investis sur le terrain, la fin de la période d'expérimentation a signifié la fin des financements alloués aux projets, sans pour autant coïncider avec la date de diffusion des résultats. Pour l'évaluateur cela a des conséquences directes : pour certains programmes, une fois l'expérimentation terminée, le CRÉDOC n'a plus eu d'interlocuteurs du côté des porteurs de projet (fin des CDD prévus pour l'accompagnement et financés par le FEJ).

I.6. Un positionnement clair de l'évaluateur : une distance suffisante pour porter un jugement évaluatif

La position de l'évaluateur est délicate au sein d'une expérimentation. Une difficulté rencontrée à chaque étape de l'évaluation et, en particulier, au moment de porter un jugement évaluatif, est celle de maintenir une distance et une neutralité dans l'élaboration des conclusions.

Le CEREQ qui a évalué de nombreuses expérimentations pour la jeunesse note bien l'impossible position de neutralité du sociologue dans le cadre de l'expérimentation[39] : le métier de sociologue consiste à « opérer une rupture par rapport au sens commun et mettre en évidence les présupposés tacites de ses sujets d'étude ». Or le guide méthodologique demandait à l'évaluateur de rester dans la logique des projets et de ne pas questionner la problématique de l'expérimentateur. Sa tâche consistant à vérifier l'impact des projets. Il est impossible d'interroger les prénotions, de faire le travail de construction de l'objet spécifique à la démarche du sociologue et d'adopter une posture critique.

Comment disposer alors de la distance nécessaire à la réflexion méthodologique ?

En participant à des comités de pilotage en observateur, le CRÉDOC a pu suivre le déroulement des expérimentations sans intervenir sur leur contenu. Il a aussi pu s'assurer de la bonne mise en œuvre des outils de suivi tels que les tableaux de bord.

---

[38] Les porteurs ont également rédigé des rapports à destination de la MAFEJ.

[39] Divay, 2012.

II. VALIDITE DES RESULTATS ET CONDITIONS DE GENERALISATION

II.1. Une durée minimale d'expérimentation pour des résultats à court et moyen terme

Dans le cadre du Fonds d'expérimentation pour la Jeunesse, deux à trois années étaient dévolues au déroulement de l'expérimentation et de l'évaluation. Concrètement, cette durée recouvrait entre autres le lancement du projet et sa montée en charge comprenant parfois la mise en place des outils de suivi. En pratique, les évaluations ont été menées sur deux ans environ. Dans plusieurs cas, cette durée s'est révélée trop courte pour mener à bien l'évaluation avec suffisamment de recul, notamment quand les projets n'étaient pas complètement « ficelés » au moment du démarrage. Certains projets se sont en outre sans doute montrés trop ambitieux à l'instar d'un projet d'accompagnement dans l'alternance. Etant donné que la majorité des contrats en apprentissage se déroulent sur deux années, le recul était insuffisant pour évaluer les taux d'obtention du diplôme et l'accompagnement dans l'emploi en vue de sécuriser le parcours des jeunes.

Tout en calibrant bien les phases de faisabilité et de montée en charge, il convient de s'assurer d'avoir un recul minimal pour évaluer les effets d'un dispositif. Cette durée dépend du contenu des actions expérimentées.

De façon plus générale, les résultats des évaluations sont à interpréter comme des résultats à court ou moyen terme. Il ne fait aucun doute que l'insertion socioprofessionnelle des jeunes se construit sur plusieurs années, et est d'autant plus longue qu'ils sortent du système scolaire avec un faible niveau de qualification ou ont connu des périodes de détention.

Les protocoles d'évaluation mis en œuvre ont porté sur l'impact des projets sur une période de deux à trois ans et ne permettent pas de tirer des conclusions sur les effets à plus long terme des dispositifs mis en œuvre.

Les retards dans la montée en charge des dispositifs ont parfois aussi réduit la période d'observation. Il est probable que dans certains cas où aucun effet n'a pu être mesuré, cela soit en réalité dû au fait qu'on ne disposait pas d'un recul suffisant pour observer des effets significatifs. De même, la mise en place de partenariats est souvent longue ; les conséquences du rapprochement entre différents acteurs peuvent n'être effectives qu'après la fin de la période d'expérimentation.

Ainsi, les effets des expérimentations évaluées sont le plus souvent limités dans le temps. Cette limite elle mérite d'être soulignée car elle doit être présente à l'esprit non seulement des évaluateurs chargés de formuler les résultats, mais aussi des utilisateurs de ces résultats. Il ne

s'agit pas de renoncer à toute velléité d'évaluation, mais simplement d'attirer l'attention des porteurs de projet et des évaluateurs sur l'importance d'ajuster leurs objectifs comme leurs attentes à ce que la durée de l'expérimentation permet de conclure. De la même manière, les tests pharmaceutiques menés sur des patients ne permettent pas de présumer des effets à long terme de l'usage d'un médicament.

II.2. La validité externe : des effets reproductibles hors du cadre expérimental ?

La question de la validité externe des résultats d'évaluation est centrale dans l'interprétation des résultats d'une expérimentation qui a vocation, par nature, à fournir des enseignements en vue de décider d'une éventuelle généralisation du projet (extension à l'échelle nationale) ou d'un essaimage potentiel (réplication sur d'autres territoires)[40]. Autrement dit la question de la validité externe est celle de savoir si les mêmes effets se reproduiront si le projet était répliqué ailleurs. Ou encore : les effets seraient-ils les mêmes si le projet était répliqué à plus grande échelle ? C'est une des questions les plus délicates de l'évaluation car il s'agit de faire la part, dans l'interprétation des résultats, entre ce qui est lié aux caractéristiques du contexte géographique, institutionnel et économique, et ce qui relève d'un « effet propre » du dispositif. Les méthodes d'évaluation randomisées sont censées le permettre mais, comme cela a été souligné, dans la plupart des cas, elles sont très difficiles à mettre en œuvre rigoureusement. L'évaluateur doit en outre être attentif à différents facteurs qui peuvent avoir une influence sur les résultats de l'évaluation et éventuellement conduire à des biais.

II.2.1. Effets propres au caractère expérimental

De par leur caractère expérimental, les projets évalués se déroulent parfois dans des conditions nouvelles ou un contexte temporairement modifié. Deux exemples tirés de notre expérience illustrent les effets indirects que peuvent induire certains dispositifs dans le cadre d'une expérimentation.

Saturation temporaire des dispositifs

Les expérimentations en faveur de l'insertion socioprofessionnelle de jeunes sous main de justice que nous avons évaluées visaient à développer les alternatives à la détention et ont favorisé les aménagements de peine. En effet, les demandes d'aménagement ont été plus nombreuses qu'habituellement dans les maisons d'arrêt concernées (dans l'un des projets, l'aménagement de peine était un critère préalable d'éligibilité). En outre, selon les juges d'application des peines, les dossiers ont été de meilleure qualité grâce au diagnostic plus précis porté sur les besoins et les attentes des jeunes. Les dossiers présentés étaient également plus solides car le suivi post-

[40] La question de la validité interne de l'évaluation consiste à déterminer si les effets (ou l'absence d'effet) mis en évidence par l'évaluation s'interprètent bien comme l'impact du dispositif à l'intérieur de l'expérimentation.

parcours prévu dans le cadre de ces expérimentations était un facteur de sécurisation des trajectoires. Par ailleurs, l'implication des juges d'application des peines, notamment dans les projets où un aménagement de peine faisait partie intégrante du programme, a nettement favorisé l'obtention de ces aménagements car les juges avaient connaissance du suivi assuré dans le cadre des programmes expérimentés[41]. Dans l'une des maisons d'arrêt, le grand nombre de demandes d'aménagement de peine acceptées pour les jeunes faisant partie de l'expérimentation a, par conséquent, vraisemblablement retardé les possibilités d'aménagement de peine pour d'autres détenus, le nombre de dossiers examinés en commission étant limité. Cet effet d'éviction, non attendu, est plutôt positif pour les jeunes bénéficiaires du programme, mais est négatif pour les autres détenus. Il est probable que sans la contrainte de faire entrer rapidement un grand nombre de jeunes dans le dispositif (en vue de l'évaluer), ce type de situation ne se serait pas produit si les dossiers de jeunes étaient présentés progressivement en débat contradictoire. Autre exemple, l'augmentation des aménagements de peine pour les jeunes a conduit à en accueillir un plus grand nombre dans les quartiers de semi-liberté d'une maison d'arrêt ; selon les juges d'application des peines, cette concentration a conduit à des problèmes de comportements qui se sont répercutés dans les activités proposées à l'extérieur dans le cadre de l'expérimentation. Les adultes (plus de 25 ans), en moins grand nombre, n'auraient pas pu jouer le rôle de régulateurs qu'ils tiennent habituellement dans ces quartiers. Cette situation, qui ne s'est produite que dans un des quatre projets de ce type, ne s'observerait probablement pas en « rythme de croisière » où les jeunes arrivent et sont présents en quartier de semi-liberté en moins grand nombre[42].

Des affectations temporaires qui peuvent précariser les porteurs

Les financements des expérimentations ont généralement servi à financer des postes supplémentaires, notamment dans les Missions Locales qui ont pu embaucher par exemple des conseillers référents justice supplémentaires. Comme cela a été souligné dans la première partie, l'affectation de personnel sur des missions financées uniquement durant le temps de l'expérimentation peut conduire à des difficultés de recrutement.

Par ailleurs, le caractère temporaire du financement des projets peut décourager des partenaires d'orienter des personnes éligibles vers des projets qui vont s'interrompre. Existe aussi le risque que cela diminue la visibilité des projets. Dans une des expérimentations en faveur de l'emploi, une importante campagne de communication avait été déployée au départ pour faire connaître le dispositif d'accompagnement en faveur de demandeurs d'emploi. Faute de moyens, la communication s'est arrêtée et les partenaires, croyant que le dispositif n'existait plus, n'ont plus orienté de chômeurs vers ces actions, ce qui a mis temporairement en péril le projet.

---

[41] Aldeghi, Jauneau, Muller, 2012.

[42] En 2011, les jeunes de 18 à 24 ans représentent 25% de la population sous écrou (Source : ministère de la Justice).

Les effets de la motivation des porteurs de projet et des bénéficiaires

Il ne fait pas de doute que la majeure partie des porteurs de projets s'étant mobilisés pour proposer une expérimentation sont particulièrement motivés par le dispositif et en connaissent les tenants et les aboutissants. Il est possible qu'il soit plus difficile de mettre en œuvre le même projet dans un autre contexte, avec la collaboration d'acteurs qui n'en sont pas à l'origine. Des freins ou des oppositions peuvent apparaître, limitant alors l'efficacité des actions mises en place.

En cas de transposition ou de généralisation d'un dispositif, il convient ainsi de s'assurer de l'appropriation des enjeux et des objectifs par les nouveaux acteurs.

Il peut aussi arriver que la communication autour d'une expérimentation rende plus assidue la participation des bénéficiaires aux actions expérimentées, quand ils sont conscients des enjeux pour les futurs bénéficiaires en cas de pérennisation. Cela a été le cas dans un dispositif en direction des jeunes sous main de justice et qui bénéficiait d'un portage très fort par les missions locales et le SPIP. Il est également difficile d'évaluer dans quelle mesure les résultats ne sont pas affectés d'un « effet Hawthorne », c'est-à-dire ne seraient pas dus à l'expérimentation même mais au fait que les sujets ont conscience de participer à une expérience dans laquelle ils sont testés, ce qui entrainerait une plus grande motivation[43]. En outre, on ne sait toutefois pas si de tels effets sont susceptibles de se maintenir sur le long terme.

De façon plus générale, la question de la visibilité de l'expérimentation auprès des bénéficiaires se pose. Notre expérience ne nous conduit pas à formuler de préconisation en la matière sauf quand l'expérimentation comporte un échantillon témoin, auquel cas, pour les raisons évoquées en partie 2, il semble préférable de ne pas mentionner l'existence d'un service « plus » réservé à un groupe de bénéficiaires. En effet, une plus forte motivation lié à l'aspect valorisant de bénéficier d'une expérimentation peut avoir un effet d'entrainement positif sur l'ensemble du projet qui ne persisterait pas en cas de généralisation. Il est possible également que dans certains cas des bénéficiaires puissent se sentir instrumentalisés et ne pas apprécier de servir de cobayes. Globalement, il est néanmoins probable que les bénéficiaires ne fassent pas de distinction entre le service « plus » et le droit commun, même si on le leur mentionne explicitement.

II.2.1. Des interrogations sur l'extrapolation des résultats

Des effets différents par sous-population ?

Avant d'envisager la généralisation d'un dispositif, il est également important de pouvoir en mesurer les effets sur différentes sous-populations. Dans le cas contraire on peut être amené à en sous-estimer ou surestimer les effets. De telles analyses nécessitent cependant de disposer d'effectifs suffisants pour pouvoir étudier des sous-populations.

---

[43] Le terme d'« effet Hawthorne » fait référence à des expériences en psychosociologie du travail, conduites dans les années 1920-1930 dans l'usine Hawthorne à Chicago où était fabriqué du matériel électrique. Ces expériences visaient à faire varier les conditions de travail d'un groupe d'ouvriers et à comparer leur productivité avec celle d'un groupe témoin. Quels que soient les changements opérés, la productivité était augmentée. L'interprétation d'un effet Hawthorne attribue cette amélioration de la productivité au fait de participer à une expérimentation.

On observe fréquemment que les effets d'un dispositif ne sont pas les mêmes selon le profil des bénéficiaires. Ainsi, les actions de formations à destination des jeunes sortant de prison ont été perçues comme très utiles et ont contribué à restaurer la confiance en soi des jeunes qui étaient de faible niveau scolaire ou très éloignés de l'emploi, tandis qu'elles ont été jugées moins pertinentes pour ceux qui avaient déjà une expérience professionnelle.

Des solutions qui fonctionnent à petite échelle mais pas dans un cadre généralisé

La particularité des expérimentations pour la jeunesse est la petite taille des projets évalués, souvent mis en place à l'échelle d'une agglomération. Des accompagnements fortement personnalisés ont souvent permis de trouver des solutions de logement, d'emploi ou de stages par exemple, pour les bénéficiaires entrés dans les dispositifs et présents en petit nombre. Il est probable que les mêmes solutions ne soient pas efficaces à plus grande échelle, s'il s'agissait de trouver un logement ou un emploi à un plus grand nombre de jeunes. Quand bien même le dispositif est totalement pertinent, les effets escomptés peuvent se heurter à une offre limitée en matière d'emploi ou de logement.

Les effets systémiques

Certains dispositifs peuvent avoir des « effets systémiques » dans leur généralisation. Par exemple : un accès accru aux aménagements de peine qui, avec la préparation et l'accompagnement adéquat, semble favoriser l'insertion des jeunes sous main de justice, nécessiterait le développement de quartiers de semi-liberté, le recrutement d'un plus grand nombre de gardiens, avec le risque qu'ils soient potentiellement moins bien formés, ce qui contrebalancerait l'effet positif de l'aménagement de peine. Autre exemple, en lien avec les limites relevées dans le paragraphe précédent : une expérimentation pour améliorer les qualifications peut produire des effets positifs sur les parcours de jeunes ciblés, mais si tous les jeunes sans qualification en acquerraient une, il n'est pas dit que le nombre de postes qualifiés disponibles serait suffisant pour qu'ils obtiennent tous un emploi adapté à leurs compétences.

Difficultés pour évaluer un projet qui regroupe plusieurs zones très petites

Plusieurs projets évalués par le CRÉDOC étaient en fait des programmes qui ont été déployés dans six ou sept territoires par un regroupement de porteurs ou par un porteur au niveau national. Si en théorie, ces projets implantés dans plusieurs zones auraient pu se prêter à la comparaison des contextes locaux et conduire à des conclusions devançant la question de l'essaimage, en pratique de telles comparaisons n'étaient pas réellement possibles, notamment lorsque chaque territoire comptait un petit nombre de bénéficiaires. En outre, ces comparaisons entre territoires ne faisaient pas partie des objectifs de l'expérimentation.

La mise en œuvre de ces projets, et le suivi des bénéficiaires, a en fait été réalisée de façon assez distincte d'un territoire à l'autre. Outre que les tensions sur le marché du logement étaient plus fortes dans certaines zones et les partenariats étaient beaucoup plus développés dans certaines

autres, le public cible n'a pas été défini de la même façon ; l'accompagnement ne s'est pas déroulé à l'identique ; malgré la définition d'un tableau de bord commun, les usages n'en ont pas été les mêmes. En particulier les rubriques n'ont pas été utilisées et remplies de façon homogène. Sans aucun doute, certains enseignements globaux ont néanmoins pu être tirés de ces différences. Toutefois, la conception d'une seule et unique expérimentation sur plusieurs territoires très différents n'est fertile que si les acteurs ont une culture commune et des pratiques harmonisées, c'est-à-dire des outils partagés et connus de tous, et surtout un pilotage fort. Les tableaux de bord ont par exemple été facilement mobilisables et interprétables dans un projet qui associait plusieurs missions locales car elles étaient étroitement coordonnées par l'une d'elles. En revanche, si les sites sont petits, les contextes socio-économiques trop variés et les pratiques de suivi des bénéficiaires hétérogènes d'un site à l'autre, une évaluation même harmonisée ne pourra pas permettre de tirer des enseignements généraux.

II.2.2. La force des expérimentations

Il reste que certaines expérimentations ont une validité externe d'une intensité variable mais cela n'enlève rien à leur force ni à leur utilité. Pour le comprendre, il faut absolument « renoncer à un usage naïf de l'évaluation »[44] qui livrerait des résultats univoques et immédiatement transposables. L'évaluation doit en effet permettre d'« identifier et faire connaître les difficultés à contourner et les démarches qui permettront de la décliner ailleurs »[45], pas forcément de manière à la reproduire à l'identique.

La trame fournie par la MAFEJ pour la rédaction des rapports d'évaluation finaux rappelle en effet qu'au-delà de la question de la validité externe des résultats de l'évaluation prise au sens strict, « les évaluateurs seront attentifs à tirer des enseignements ayant une valeur générique : il ne s'agit pas seulement de se prononcer sur la valeur ajoutée du dispositif "clé en main", mais d'identifier les mécanismes à l'œuvre et d'en tirer des conclusions utiles pour la conception et la conduite des politiques publiques, y compris pour d'autres acteurs et d'autres territoires. »[46]

La généralisation d'une expérimentation se faisant dans une période ultérieure à celle de la mise en œuvre au stade expérimental, les résultats peuvent différer quand le contexte économique change entre ces deux périodes. Cette situation s'est produite dans le cas de la généralisation du Revenu de Solidarité Active qui s'est fait dans un contexte de crise économique nettement aggravée par rapport au moment de son expérimentation, réduisant d'autant les opportunités d'insertion professionnelle durable des allocataires.

---

[44] Gurgand, Valdenaire, 2012.

[45] *Idem*.

[46] *Idem*.

III. USAGES ET UTILITE DES EVALUATIONS

III.1. Usage et utilité des évaluations par les porteurs

Si, pour les évaluateurs, les expérimentations ont représenté une expérience riche d'enseignements méthodologiques et une opportunité d'étendre leurs recherches dans le champ de la jeunesse, les porteurs de projet ont eu, quant à eux, une approche pragmatique en lien avec leur métier de terrain.

Pour les porteurs de projets et leurs partenaires, les expérimentations financées par le FEJ ont incontestablement représenté une opportunité d'élargir leur champ d'action, de développer des partenariats locaux, de renforcer les actions existantes, de tester des actions nouvelles ou parfois simplement de faire bénéficier à un public plus large les actions déjà existantes. L'intérêt scientifique des expérimentations sociales est parfois éloigné des préoccupations des professionnels qui sont confrontés quasi quotidiennement à des publics en difficulté et parfois en détresse. Pour nombre d'entre eux toutefois, l'évaluation, si elle a pu apparaître initialement comme une contrainte, devait finalement apporter les preuves objectives de l'utilité et l'efficacité de leur action pour pouvoir obtenir des financements. Aussi l'ont-ils acceptée et ont-ils formulé de fortes attentes envers ses résultats.

Il est toutefois probable que la plupart d'entre eux se sont lancés dans ces projets, forts de la conviction que les actions expérimentées étaient pertinentes et qu'ils trouveraient un moyen de les maintenir, quels que soient les résultats des évaluations. Le FEJ apportait des financements à hauteur de 50% maximum du budget des projets, financements qui s'interrompaient à la fin des deux à trois ans d'expérimentation. Ce cadre a généré un sentiment d'incompréhension de la part de beaucoup de porteurs de projet. En effet, s'ils désiraient prolonger leurs actions, ce qui était le souhait de la plupart d'entre eux, ils devaient donc se mobiliser dans la recherche de financements avant la fin de la contribution du FEJ, et par conséquent bien avant l'établissement des résultats finaux de l'évaluateur « officiel ». Ainsi, un centre communal d'action sociale, par exemple, a obtenu avant la fin de l'expérimentation un financement direct pour mettre en place le dispositif de microcrédit qu'il a « expérimenté » avec le soutien financier du FEJ (CCAS Besançon). De même, une plate-forme pour l'orientation vers et l'accompagnement pendant l'apprentissage a été pérennisée bien avant la fin de l'évaluation grâce à un mécénat et des financements de la politique de la ville (Fondation d'Auteuil). Dans d'autres cas, quelques actions ont été maintenues. Un dispositif d'insertion de jeunes sortant de prison a été prolongé pour deux ans avec le soutien du Préfet à l'Égalité des chances et de collectivités territoriales. Enfin, un projet est en voie de pérennisation et pour divers autres projets, des recherches de financement sont actuellement en cours.

Paradoxalement, les porteurs de projet n'ont pas pour autant rejeté le principe de l'évaluation. Anticipant le fait que les résultats ne seraient disponibles que bien après la fin des financements, plusieurs porteurs ont en fait mis en place leurs propres évaluations en interne - devançant donc parfois celles du CRÉDOC. Tous n'ont pas eu besoin de le faire pour trouver des financements auprès de partenaires locaux qui avaient pu constater localement des effets positifs. Mais les

évaluations menées en parallèle par les évaluateurs extérieurs ont sans doute jeté le discrédit sur certains résultats qui arrivaient « trop tard » et ont pu nuire à nos relations avec les porteurs. On peut craindre qu'ils ne soient pas prêts à recommencer une expérimentation : savoir que les financements sont temporaires n'incitera pas forcément des porteurs à mobiliser leurs forces et monter des projets qui ont une durée de vie déterminée.

Ces constats posent évidemment la question de l'utilité, pour les porteurs, des évaluations financées par le FEJ. On peut se réjouir que cela favorise le développement d'une culture de l'évaluation en interne, sensibilise les acteurs publics et privés aux méthodes quantitatives et qualitatives des chercheurs, et leur donnent des clés d'aide à la décision. Restent toutefois en suspens la question de l'utilité de ce vaste chantier de petites expérimentations et la question du décalage entre la temporalité de l'action de terrain et celle de l'action politique. Jusqu'à présent, ces expérimentations lancées par le FEJ n'ont pas débouché sur des politiques nationales en faveur de la jeunesse, le contexte budgétaire s'étant révélé peu favorable.

III.2. Enseignements de politiques publiques

L'un des intérêts des expérimentations n'est pas d'aboutir à la pérennisation ou la généralisation de l'ensemble de ces dispositifs mais de faire émerger des idées, des bonnes pratiques. En réalité, l'usage qui est fait des expérimentations implique une « capacité d'extrapolation »[47].

La thématique du logement, abordée sous des formes très diverses dans les projets, s'est révélée très complexe car le logement est un élément central du processus d'autonomisation des jeunes qui dépend étroitement du parcours personnel et professionnel. De plus, la question du logement est naturellement fortement liée au contexte local. Enfin, elle ne se prête probablement pas à un suivi à long terme.

Les plateformes de lutte contre le décrochage ont été généralisées avant la fin des expérimentations, renforçant les liens partenariaux tissés en cours d'expérimentation.

Si les porteurs de projets ont pu se montrer impatients ou déçus vis-à-vis des résultats de l'évaluation, il est incontestable que les expérimentations ont conduit à de nombreux échanges et à de riches partages d'idées, en particulier lors des journées thématiques organisées tous les 6 mois sur chaque thème par la MAFEJ. D'après l'expérience du CRÉDOC, la MAFEJ n'a pas piloté de façon rapprochée l'ensemble des projets, leur nombre très élevé empêchant sans doute une forte implication de sa part. Les porteurs ont pu s'emparer des résultats de projets proches des leurs par le biais des réseaux des CCAS ou ceux des conseils généraux, ou à travers le rôle du Conseil National des Missions Locales.

Le travail de capitalisation des résultats de cet ensemble très riche d'expérimentations reste à réaliser. Pour l'instant aucun enseignement global n'a été tiré officiellement de tous ces projets.

---

[47] Gurgand, Valdenaire, 2012.

CONCLUSION

Les expérimentations financées par le Fonds d'expérimentation pour la Jeunesse (FEJ) ont couvert une grande diversité d'actions. Certains dispositifs ont par exemple proposé un accompagnement renforcé des bénéficiaires par rapport à celui existant dans le droit commun, tandis que d'autres ont expérimenté des actions nouvelles (par exemple des ateliers et des formations spécifiques à destination des jeunes sous main de justice, dans le cadre d'un aménagement de peine) ou élargi aux jeunes leur public cible. Au sein de chaque projet, les objectifs fixés par l'évaluation étaient multiples : mesure de l'impact sur les bénéficiaires en termes d'accès au logement, d'accès à l'emploi, de parcours dans l'apprentissage, de récidive, de processus d'autonomie, d'estime de soi, analyse des conditions de mise en œuvre, mesure de l'impact sur les acteurs des dispositifs et les partenariats.

Les évaluations menées par le CRÉDOC depuis 2009 ont permis de mettre en évidence certains effets sur les bénéficiaires et de tirer des enseignements sur les conditions de mise en œuvre des projets. Les évaluations conduites ont principalement permis de porter un jugement évaluatif sur les aspects *qualitatifs* des projets, c'est-à-dire les conditions de mise en œuvre des actions et les dynamiques partenariales qui les sous-tendent. Les mesures d'impact des projets sur les bénéficiaires ont cependant été limitées en raison des difficultés rencontrées pour mettre en place des échantillons témoins aléatoires et suivre les bénéficiaires et témoins dans le temps. Le principal obstacle au déploiement des méthodes quantitatives est en définitive lié aux effectifs réduits de jeunes concernés par les expérimentations évaluées.

L'expérience du CRÉDOC rejoint celle d'autres chercheurs, notamment le CEREQ, qui a également évalué un nombre important d'expérimentations dans le cadre du FEJ. Les constats dressés remettent en question le dogme de la randomisation. L'expérimentation basée sur l'assignation aléatoire des potentiels bénéficiaires à un groupe test ou à un groupe témoin est aujourd'hui toujours érigée en règle d'or[48] et est la méthode préconisée dans les appels à projets de la Commission Européenne. Difficile à mettre en œuvre en pratique, elle l'est tout particulièrement dans le cadre de projets de faible envergure qui concernent une population de taille réduite. En termes de méthodes, notre bilan nous incite à ne pas préconiser une utilisation exclusive des approches randomisées. L'approche qualitative a en effet permis d'apporter des informations précises et précieuses pour l'évaluation, notamment sur les questions de mise en œuvre des dispositifs et de fonctionnement des partenariats. Nous recommandons ainsi à l'avenir de développer davantage d'outils qualitatifs : observations sur site, entretiens approfondis de type « récits de vie »… De façon plus générale, il importe de ne pas perdre de vue les autres méthodes d'évaluation et ne pas oublier que l'expérimentation reste une forme d'évaluation parmi d'autres.

---

[48] Le terme de « gold standard » a été mentionné en 1993 par les économistes Rossi et Freeman au sujet de la randomisation (Rossi, Freeman, 1993).

D’autres facteurs ont par ailleurs conditionné la nature et la portée des résultats produits : le ciblage plus ou moins précis des populations éligibles, le caractère plus ou moins innovant des actions expérimentées ou encore la durée de l’évaluation souvent amputée par une montée en charge trop longue. A ce titre, la réalisation d’études de faisabilité et la mise en place de phases pilotes préalables à l’expérimentation, font partie des principales préconisations que nous formulons.

Aujourd’hui, parmi les évaluateurs comme au sein de la MAFEJ, le mot d’ordre est la « capitalisation ». Sans cela quel sens donner à autant de petits projets ? Quel apport pour les politiques en faveur de la jeunesse ? Indubitablement, les centaines de projets expérimentés et évalués dans le cadre du Fonds d’expérimentation pour la Jeunesse constituent un vivier d’idées dans lequel il sera possible de puiser àl’avenir. Les porteurs de projets et leur partenaires n’ont généralement pas attendu les résultats de l’évaluation, ni même la fin des expérimentations pour prolonger leurs actions. Les échanges entre porteurs sont nombreux, que ce soit sous l’égide de la MAFEJ et ses ateliers thématiques ou à travers les réseaux tels que le CNML ou l’UNCASS. Sans doute, le Fonds d’expérimentation pour la Jeunesse aura-t-il joué un rôle de catalyseur des politiques locales de le jeunesse. L’important aujourd’hui est de prolonger le partage de bonnes pratiques afin de continuer à dynamiser les concepteurs et acteurs des politiques sociales. La MAFEJ vient de commencer le travail de synthèse des centaines de projets financés, dont la plupart se terminent tout juste actuellement. Pour les évaluateurs que nous sommes, le travail se poursuit donc sur le terrain de la réflexion méthodologique et des échanges entre chercheurs, qui pourront s’envisager au travers de réseaux d’experts au niveau national, mais aussi au niveau européen.

BIBLIOGRAPHIE

Isa ALDEGHI, Pauline JAUNEAU et Lara MULLER, « Quatre expérimentations pour favoriser l'insertion des jeunes sous main de justice », *Consommation et mode de vie* n°256, CRÉDOC, octobre 2012.

Guillaume ALLEGRE, L'expérimentation du RSA et ses ambiguïtés, Informations sociales n°174, CNAF, novembre-décembre 2012, pp.52-60.

Denis ANNE et Yannick L'HORTY, « Aides sociales locales, revenu de Solidarité active (RSA) et gains du retour à l'emploi », *Economie et Statistique* n°429-430, INSEE, 2009, pp. 129-157.

Jean-Claude BARBIER, « Évaluation "expérimentales" : quelques leçons de l'histoire », *Politiques sociales et familiales* n°110, CNAF, décembre 2012, pp.19-31.

Jean-Claude BARBIER et Nicolas MATY-JASIK, « Évaluation des politiques publiques et quantification en France : des relations ambiguës et contradictoires entre disciplines », *Revue Française de Sciences Economiques* n°5, 2010, pp. 123-140.

Luc BEHAGHEL, Bruno CREPON et Marc GURGAND, *Evaluation d'impact de l'accompagnement des demandeurs d'emploi par les opérateurs privés de placement et le programme Cap vers l'entreprise - Rapport final*, CREST-Ecole d'Economie de Paris, Septembre 2009, 59 pages.

Manon BREZAULT, Christine OLM, Matthieu ANGOTTI et Marie-Odile SIMON, *Politiques locales de la jeunesse et expérimentations sociales – Vers une cohérence des politiques de la jeunesse ?* CRÉDOC, Cahier de recherche n°266, décembre 2009.

Elisa CHELLE, « Expérimentation sociale : la tentation américaine », *Informations sociales* n°174, CNAF, novembre-décembre 2012, pp.24-30.

Julien DAMON, « La fièvre de l'évaluation », *Sciences Humaines* n°208, octobre 2009.

Sophie DIVAY, « L'évaluation qualitative : objectiver sans compter », *Bref du CEREQ* n°286, CEREQ, mai 2011.

Sophie DIVAY, « "Sociologue-évaluateur" : un oxymore ? » *Informations sociales* n°174, CNAF, novembre-décembre 2012, pp.114-120.

Pauline DOMINGO, « Évaluation des effets de l'accompagnement sur les trajectoires des bénéficiaires : quels enseignements ? », *Informations sociales* n° 169, CNAF, 2012, pp. 100-107.

Fonds d'expérimentation pour la Jeunesse, *Guide méthodologique pour l'évaluation des expérimentations sociales à l'intention des porteurs de projets*, avril 2009, 22 pages.

Fonds d'expérimentation pour la Jeunesse, *Rapport d'activité 2009-2011*, (sans date), 48 pages.

Fonds d'expérimentation pour la Jeunesse, *Rapport du Conseil scientifique du Fonds d'expérimentation pour la jeunesse pour l'année 2011*, 19 pages.

Pauline GIVORD, « Méthodes économétriques pour l'évaluation de politiques publiques », Document de Travail N° G2010/08, Insee.

Marc GURGAND et Mathieu VALDENAIRE, « Expérimentations : conception et pilotage - Le Fonds d'expérimentation pour la jeunesse et les politiques éducatives : premier retour d'expérience », *Éducation & formations* n° 81, mars 2012, pp. 27-37.

Annie FOUQUET, « L'usage des statistiques : de l'aide à la décision à l'évaluation des politiques publiques », *Revue Française des Affaires sociales* n°1-2, 2010, pp.309-322.

Léopold GILLES, « Les évaluations conduites dans le cadre du Fonds d'expérimentation pour la Jeunesse : premier retour d'expérience », *Informations sociales* n°174, CNAF, novembre-décembre 2012, pp.66-73.

Bernard GOMEL et Évelyne SERVERIN, *Évaluer l'expérimentation sociale,* Centre d'Études de l'Emploi, *Document de travail* n°143, 21 pages.

J-Pal Europe, *Social experimentation – A methodological guide for policy makers*, rédigé à la demande de la Direction Générale de l'Emploi, des Affaires sociales et de l'Inclusion de la Commission européenne, septembre 2011, 30 pages.
http://ec.europa.eu/social/BlobServlet?docId=7102&langId=en

Marjorie JOUEN, « L'expérimentation sociale prend sa place dans l'Union européenne », *Informations sociales* n°174, CNAF, novembre-décembre 2012, pp.32-40.

Agnès LABROUSSE, « Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective d'un outil de preuve et de gouvernement », *Revue de la régulation* n°7, 1er semestre 2010, pp.2-30.

Philippe LEMISTRE, « L'évaluation en sciences sociales est-elle un mythe? Le cas des politiques éducatives », in *Evaluation et données longitudinales : quelles relations* ? Actes des XVIIèmes journées d'étude sur les données longitudinales dans l'analyse du marché du travail, Marseille, 27-28 mai 2010, Relief n° 30 , 2010 , pp.111-121.

François LEGENDRE, « Le recours croissant aux expérimentations sociales en économie », *Informations sociales* n°174, CNAF, novembre-décembre 2012, pp.14-22.

Peter H. ROSSI et Howard E. FREEMAN, « Evaluation : a Systematic Approach », Newbury Park, Sage, 1993.

Patrick VIVERET, *L'évaluation des politiques et des actions publiques : Propositions en vue de l'évaluation du Revenu Minimum d'Insertion* – Rapport au Premier Ministre, Observatoire de la décision publique, juin 1989, 57 pages.